Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

Lecture Note On Mathematical Statistics 1

B.Sc. in Mathematics

Fourth Stage

Assist. Prof. Dr. Feras Shaker Mahmood

Syllabus of Mathematical Statistics 1

- **Chapter 1: Additional Topics in Probability**
- **Special Distribution Functions : The Binomial Probability Distribution , Poisson Probability Distribution , Uniform Probability Distribution , Normal Probability Distribution , Gamma Probability Distribution , Distributions of Functions of random Variables (Transformation technique, Distribution Function technique, Moment generating function technique), Limit Theorems: Chebyshev`s Theorem Law of Large Numbers, Central Limit Theorem.**
- **Chapter 2: Sampling Distributions**
- **Sampling Distributions Associated with Normal Populations, Distribution of $\overline{X}$ and S^2 , Chi-Square Distribution, Student t-Distribution, F-Distribution, Distributions of Order statistics, Large sample Approximations: The Normal Approximation to the Binomial Distribution, Limiting Distribution: Stochastic Convergence, Limiting of moment generating functions, Theorems on Limiting distributions.**
- **Chapter 3: Point Estimation**
- **The Method of Moments, The Method of Maximum Likelihood, Some desirable properties of point estimators, Unbiased Estimators, Sufficiency, Consistency, Efficiency, Minimal Sufficiency and Minimum-Variance Unbiased Estimation, Cramer–Rao procedure to test for efficiency.**

References

- الاحصاء الرياضي، امير حنا ، الطبعة الاولى، جامعه الموصل، ١٩٩٥.
- مقدمة في الاحصاء الرياضي، صباح داود سليم، الطبعة الاولى، جامعة البصرة، ١٩٨٩.
- **Introduction to Mathematical Statistics, R. V. Hogg and A.T. Craig, (4, 5, 6) edition.**
- **Introduction to the Theory of Statistics, Alexander M.Mood, Franklin, A. Graybill, MC- Graw Hill, 1974.**
- **Mathematical Statistics with Applications, K. M. Ramachandran and C. P. Tsokos, 2009.**
- **Mathematical Statistics with Application, Richard L.Schaffer, Seventh Edition, Thomson Brooks, 2008.**
- **Probability and mathematical Statistics, Prasanna Sahoo, University of Louisville,, USA, 2008.**
- **An introduction to Mathematical Statistics and Its Applications, Ricard I. larson and Morris Luois, Fifth edition, Person Education, INC, 2012.**
- **Probability and Statistical Inference, Robert V. Hogg ,Elliot A. Tanis and ale L. Zimmerman, Ninth Edition, Pearson Education, USA,2015.**

Discrete Distributions

Bernoulli
$0 < p < 1$

$$f(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1$$
$$M(t) = 1 - p + pe^t, \qquad -\infty < t < \infty$$
$$\mu = p, \qquad \sigma^2 = p(1-p)$$

Binomial
$b(n, p)$
$0 < p < 1$

$$f(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}, \qquad x = 0, 1, 2, \ldots, n$$
$$M(t) = (1 - p + pe^t)^n, \qquad -\infty < t < \infty$$
$$\mu = np, \qquad \sigma^2 = np(1-p)$$

Geometric
$0 < p < 1$

$$f(x) = (1-p)^{x-1}p, \qquad x = 1, 2, 3, \ldots$$
$$M(t) = \frac{pe^t}{1-(1-p)e^t}, \qquad t < -\ln(1-p)$$
$$\mu = \frac{1}{p}, \qquad \sigma^2 = \frac{1-p}{p^2}$$

Hypergeometric
$N_1 > 0, \;\; N_2 > 0$
$N = N_1 + N_2$

$$f(x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}, \qquad x \le n, x \le N_1, n - x \le N_2$$
$$\mu = n\left(\frac{N_1}{N}\right), \qquad \sigma^2 = n\left(\frac{N_1}{N}\right)\left(\frac{N_2}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Negative Binomial $f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \qquad x = r, r+1, r+2, \ldots$

$0 < p < 1$

$r = 1, 2, 3, \ldots$ $M(t) = \dfrac{(pe^t)^r}{[1-(1-p)e^t]^r}, \qquad t < -\ln(1-p)$

$$\mu = r\left(\frac{1}{p}\right), \qquad \sigma^2 = \frac{r(1-p)}{p^2}$$

Poisson $f(x) = \dfrac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0, 1, 2, \ldots$

$\lambda > 0$

$$M(t) = e^{\lambda(e^t-1)}, \qquad -\infty < t < \infty$$

$$\mu = \lambda, \qquad \sigma^2 = \lambda$$

Uniform $f(x) = \dfrac{1}{m}, \qquad x = 1, 2, \ldots, m$

$m > 0$

$$\mu = \frac{m+1}{2}, \qquad \sigma^2 = \frac{m^2-1}{12}$$

Continuous Distributions

Beta
$\alpha > 0$
$\beta > 0$

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1$$

$$\mu = \frac{\alpha}{\alpha+\beta}, \qquad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

Chi-square
$\chi^2(r)$
$r = 1, 2, \ldots$

$$f(x) = \frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-x/2}, \qquad 0 < x < \infty$$

$$M(t) = \frac{1}{(1-2t)^{r/2}}, \qquad t < \frac{1}{2}$$

$$\mu = r, \qquad \sigma^2 = 2r$$

Exponential
$\theta > 0$

$$f(x) = \frac{1}{\theta} e^{-x/\theta}, \qquad 0 \le x < \infty$$

$$M(t) = \frac{1}{1-\theta t}, \qquad t < \frac{1}{\theta}$$

$$\mu = \theta, \qquad \sigma^2 = \theta^2$$

Gamma

$\alpha > 0$

$\theta > 0$

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-x/\theta}, \qquad 0 < x < \infty$$

$$M(t) = \frac{1}{(1-\theta t)^{\alpha}}, \qquad t < \frac{1}{\theta}$$

$$\mu = \alpha\theta, \qquad \sigma^2 = \alpha\theta^2$$

Normal

$N(\mu, \sigma^2)$

$-\infty < \mu < \infty$

$\sigma > 0$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \qquad -\infty < x < \infty$$

$$M(t) = e^{\mu t + \sigma^2 t^2/2}, \qquad -\infty < t < \infty$$

$$E(X) = \mu, \qquad \mathrm{Var}(X) = \sigma^2$$

Uniform

$U(a, b)$

$-\infty < a < b < \infty$

$$f(x) = \frac{1}{b-a}, \qquad a \le x \le b$$

$$M(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}, \qquad t \ne 0; \qquad M(0) = 1$$

$$\mu = \frac{a+b}{2}, \qquad \sigma^2 = \frac{(b-a)^2}{12}$$

Table X Discrete Distributions

Probability Distribution and Parameter Values	Probability Mass Function	Moment-Generating Function	Mean $E(X)$	Variance $\mathrm{Var}(X)$	Examples
Bernoulli $0 < p < 1$ $q = 1 - p$	$p^x q^{1-x}, \quad x = 0, 1$	$q + pe^t$, $-\infty < t < \infty$	p	pq	Experiment with two possible outcomes, say success and failure, $p = P(\text{success})$
Binomial $n = 1, 2, 3, \ldots$ $0 < p < 1$	$\binom{n}{x} p^x q^{n-x}$, $x = 0, 1, \ldots, n$	$(q + pe^t)^n$, $-\infty < t < \infty$	np	npq	Number of successes in a sequence of n Bernoulli trials, $p = P(\text{success})$
Geometric $0 < p < 1$ $q = 1 - p$	$q^{x-1} p$, $x = 1, 2, \ldots$	$\dfrac{pe^t}{1 - qe^t}$ $t < -\ln(1 - p)$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	The number of trials to obtain the first success in a sequence of Bernoulli trials
Hypergeometric $x \le n, x \le N_1$ $n - x \le N_2$ $N = N_1 + N_2$ $N_1 > 0, \quad N_2 > 0$	$\dfrac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$		$n\left(\dfrac{N_1}{N}\right)$	$n\left(\dfrac{N_1}{N}\right)\left(\dfrac{N_2}{N}\right)\left(\dfrac{N-n}{N-1}\right)$	Selecting n objects at random without replacement from a set composed of two types of objects
Negative Binomial $r = 1, 2, 3, \ldots$ $0 < p < 1$	$\binom{x-1}{r-1} p^r q^{x-r}$, $x = r, r+1, \ldots$	$\dfrac{(pe^t)^r}{(1 - qe^t)^r}$, $t < -\ln(1 - p)$	$\dfrac{r}{p}$	$\dfrac{rq}{p^2}$	The number of trials to obtain the rth success in a sequence of Bernoulli trials
Poisson $\lambda > 0$	$\dfrac{\lambda^x e^{-\lambda}}{x!}$, $x = 0, 1, \ldots$	$e^{\lambda(e^t - 1)}$ $-\infty < t < \infty$	λ	λ	Number of events occurring in a unit interval, events are occurring randomly at a mean rate of λ per unit interval
Uniform $m > 0$	$\dfrac{1}{m}, \quad x = 1, 2, \ldots, m$		$\dfrac{m+1}{2}$	$\dfrac{m^2 - 1}{12}$	Select an integer randomly from $1, 2, \ldots, m$

Table XI Continuous Distributions

Probability Distribution and Parameter Values	Probability Density Function	Moment-Generating Function	Mean $E(X)$	Variance $\text{Var}(X)$	Examples
Beta $\alpha > 0$ $\beta > 0$	$\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, $0 < x < 1$		$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$	$X = X_1/(X_1+X_2)$, where X_1 and X_2 have independent gamma distributions with same θ
Chi-square $r = 1, 2, \ldots$	$\dfrac{x^{r/2-1}e^{-x/2}}{\Gamma(r/2)2^{r/2}}$, $0 < x < \infty$	$\dfrac{1}{(1-2t)^{r/2}}$, $t < \dfrac{1}{2}$	r	$2r$	Gamma distribution, $\theta = 2$, $\alpha = r/2$; sum of squares of r independent $N(0,1)$ random variables
Exponential $\theta > 0$	$\dfrac{1}{\theta}e^{-x/\theta}$, $0 \le x < \infty$	$\dfrac{1}{1-\theta t}$, $t < \dfrac{1}{\theta}$	θ	θ^2	Waiting time to first arrival when observing a Poisson process with a mean rate of arrivals equal to $\lambda = 1/\theta$
Gamma $\alpha > 0$ $\theta > 0$	$\dfrac{x^{\alpha-1}e^{-x/\theta}}{\Gamma(\alpha)\theta^{\alpha}}$, $0 < x < \infty$	$\dfrac{1}{(1-\theta t)^{\alpha}}$, $t < \dfrac{1}{\theta}$	$\alpha\theta$	$\alpha\theta^2$	Waiting time to αth arrival when observing a Poisson process with a mean rate of arrivals equal to $\lambda = 1/\theta$
Normal $-\infty < \mu < \infty$ $\sigma > 0$	$\dfrac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$, $-\infty < x < \infty$	$e^{\mu t+\sigma^2 t^2/2}$ $-\infty < t < \infty$	μ	σ^2	Errors in measurements; heights of children; breaking strengths
Uniform $-\infty < a < b < \infty$	$\dfrac{1}{b-a}$, $a \le x \le b$	$\dfrac{e^{tb}-e^{ta}}{t(b-a)}$, $t \ne 0$; 1, $t = 0$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	Select a point at random from the interval $[a, b]$

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور

فراس شاكر محمود

Distribution function method

Basically the method of distribution function is as follows If x is a random variable with pdf $f_x(x)$ and if y is some function of x , then we can find the cdf $f_y(y)$ = p(Y$\leq y$) directly by integrating $f_x(x)$ over the region for which (Y$\leq y$) , now by differentiating $f_y(y)$, we get the probability density function $f_y(y)$ of Y . In general , if Y is a function of random variable x_1.....x_2 say g (x_1.....x_n), then we can summarize the method of distribution function as follows.

PROCEDURE TO FIND CDF OF A FUNCTION OF R.V USING THE METHOD OF DISTRIBUTION FUNCTIONS.

1- find the region (Y$\leq y$) in the (x_1, x_2......,x_n)space that is find the set of (x_1, x_2......,x_n)for which g((x_1,......,x_n) $\leq y$.

2- find fY(y)= p(Y$\leq y$) by integrating (x_1, x_2......,x_n) over the region (Y$\leq y$).

3- find the distribution function $f_y(y)$ by differentiating fY(y).

Example: let x~N(0,1) using the cdf of x find the pdf of y=x^2

Solution:

Note that the pdf of X is

$$f_x(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}} \qquad -\infty < x < \infty$$

then the cumulative distribution function of Y for a given y> 0 is fY(y)= p(Y$\leq y$) = p($e^x \leq y$)

$$= p(x \leq ln\ y)$$

$$= \int_{-x}^{In\ y} \frac{1}{2\pi} e^{\frac{x^2}{2}} \quad dx$$

Hence by differentiating $f_y(y)$, we obtain the probability density function as.

$$f(y)=\begin{cases} \frac{1}{y\sqrt{2\pi}}\, e^{\frac{x^2}{2}} & 0 < y \\ 0 & \text{other wise} \end{cases}$$

Example: let $f_{(x)} = \frac{1}{x^2}$, $x \geq 1$ find the p. d. f., $Y=e^{-x}$ by using distribution technique ?

Solution:

$$f(x)=\begin{cases} \frac{1}{x^2} \quad for \quad x \geq 1 \\ 1 \qquad\qquad 0.w \end{cases}$$

$$f(y) = p\left(Y \leq y\right) \Longrightarrow p\left(e^{-x} \leq y\right)$$

$$f(y) = p\left(-x \leq \ln y\right) * -1$$

Example: let x~N(0,1) using the cdf of x find the pdf of y=x^2

Solution:

Since x~N(0,1)

$$\therefore f(x)=\begin{cases} \frac{1}{\sqrt{2\,\pi}} \quad e^{\frac{-1}{2}x^2} & \text{for } -\infty < x < \infty \\ 0 & 0.w \end{cases}$$

$$f(y)=p\left(Y \leq y\right)$$

$$f(y)=p\left(x^2 \leq y\right)$$

$$f(y)=p\left(x \quad \leq \pm\sqrt{y}\right) \qquad -\sqrt{y} < x < \sqrt{y}$$

$$f(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x^2} dx$$

$\forall x^2 = 3$

$2xd = dy \Longrightarrow dx \frac{cly}{2x}$

$dx = \frac{dy}{2\sqrt{y}}$

$$f(y) = 2\frac{d}{dy}\int_{0}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{\frac{-y}{2}} \frac{1}{2\sqrt{y}} dy$$

f(y)= $\begin{cases} \frac{1}{\sqrt{2\pi}} \quad (y)^{\frac{-1}{2}} e^{\frac{-y}{2}} & 0 < y < \infty \\ 0 & \text{o.w} \end{cases}$

yn$x_{(1)}$

$-\infty < x < \infty$

$0 < x < \infty$

$0 < y < \infty$

Example: let x~N(0,1) using the c. d. f. of x. find the p. d. f. of y=e^x

Solution:

Since x~N(0,1)

f(y)= $\begin{cases} \frac{1}{\sqrt{2\pi}} \quad e^{\frac{-1}{2}} x^2 & -\infty < x < \infty \\ 0 & \text{o.w} \end{cases}$

f(y)=p$\left(Y \leq y\right) \Longrightarrow$ =p$\left(e^x \leq y\right)$

f(y)=p$\left(x \leq \ln y\right)$

$f(y) = \int_{-\infty}^{lny} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x^2} dx \Longrightarrow f(y) = \frac{d}{dy}\int_{-\infty}^{lny} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x^2} dx$

$f(y) = \frac{1}{2\pi}\ e^{\frac{-1}{2}x^2}dx$ $\qquad$ $y = e^x$

x=ln y

$$f(y) = \begin{cases} \frac{1}{y\sqrt{2\pi}} \quad e^{\frac{-1(ln\,y)^2}{2}} & 0 < y < \infty \\ 0 & \text{o. w.} \end{cases}$$

$$\infty < x < \infty$$
$$e^{\infty} < e^x > e^{\infty}$$
$$0 \leq e^x < \infty$$
$$0 < y < \infty$$

Example: If X ~ Poisson(y) find the cumulative distribution function of Y=ax+b

Solution:

Since x~N(0,1)

$$\therefore f(y) = \begin{cases} \frac{1*e^{-1}}{x_1} & \text{for} \quad x = 0 \ldots\ldots\ldots \infty \\ 0 & \text{o.w} \end{cases}$$

f(y)=p$\left(Y \leq y\right)$ $\Longrightarrow$ p$\left(ax + b \leq y\right)$

f(y)=p$\left(ax \leq y - b\right)$÷a

f(y)=p$\left(x \leq \frac{y-b}{a}\right)$

sine xupo(1) $\Longrightarrow$ discrete distribution

$$f(y) = \sum_{x=0}^{\frac{y-b}{0}} \frac{ב^x e^ב}{x1}$$

$y = ax + b$ $\qquad$ x=0,............

if x=0y = 0+b

if x=1 $\Longrightarrow$ y=a+b

if x=2 $\Longrightarrow$ y=2a+b

y=b,a+b,2a+b,3a+b,......

$\therefore y = na + b \qquad \partial, n = 0, \ldots \ldots \ldots ..$

f(y) = p $\left(x \geq -lny\right) \Longrightarrow$ f(y) = 1-p$\left(x \leq -lny\right)$

$f(y) = 1 - \int_{1}^{-1lny} \frac{1}{x^2} dx \Longrightarrow f(y) = 1 - \int_{1}^{-1lny} x^{-2}\, dx$

$f(y) = 1 - [\ \frac{1}{2}\ ]_1^{-lny} \Longrightarrow f(y) = 1 - [\ \frac{1}{lny} + 1\ ]$

$$f(y) = \frac{-1}{x}$$

$f(y) = \frac{5-(-1)\frac{1}{y}}{(lny)^2} \Longrightarrow f(y) = \frac{\frac{1}{y}}{(lny)^2}$

$f(y) = \frac{1}{y(lny)^2}$

$1 \leq \text{x} < \infty$

$-1 \geq -\text{x} > -\infty$

$\infty \leq -\text{x} < -1$

$e^{-\infty} < e^{-x} < e^{-1}$

$0 < y < e^{-1}$

$$\therefore \text{f(x)}= \begin{cases} \frac{1}{y(lny)^2} & for\ \ 0 < y < e^1 \\ 0 & \text{0.w.} \end{cases}$$

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات في مادة احصاء ١

المحاضرة الثانيه الساندة

B. Sc. in Mathematics

Second Stage

Assistant . Prof. Dr. Feras Shaker Mahmood

.

Probability Integral Transformation

Let X be a continuous random variable, with pdf f and cdf F. Let $Y = F(X)$. Then,

$$P(Y \le y) = P(F(X) \le y) = P(X \le F^{-1}(y))$$

$$= \int_{-\infty}^{F^{-1}(y)} f_X(x)dx = F_X(x)\Big|_{-\infty}^{F^{-1}(y)} = y.$$

Hence,

$$f(y) = \begin{cases} 1, & 0 < y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Thus, Y has a $U(0,1)$ distribution. The transformation $Y = F(X)$ is called a *probability integral transformation*. It is interesting to note that irrespective of the pdf of X, Y is always uniform in $(0, 1)$.

A simple generalization of the method of distribution functions to functions of more than one variable is the *transformation method*. We illustrate the method for bivariate distributions. The method is similar for the multivariate case. Let the joint pdf of (X, Y) be $f(x, y)$. Let $U = g_1(X, Y)$; $V = g_2(X, Y)$. The mapping from (X, Y) to (U, V) is assumed to be one-to-one and onto. Hence, there are functions, h_1 and h_2 such that

$$x = h_1^{-1}(u, v),$$

and

$$y = h_2^{-1}(u, v).$$

Define the Jacobian of the transformation J by

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial u} \end{vmatrix}.$$

Then the joint pdf of U and V is given by

$$f(u, v) = f(h_1^{-1}(u, v), h_2^{-1}(u, v))\,|J|\,.$$

Example

Let X and Y be independent random variables with common pdf $f(x) = e^{-x}$, $(x > 0)$. Find the joint pdf of $U = X/(X+Y)$, $V = X+Y$.

Solution

We have $U = X/(X+Y) = X/V$*. Hence,* $X = UV$ *and* $Y = V - X = V - UV = V(1-U)$*. Thus, the Jacobian*

$$J = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix}.$$

Then $|J| = v(1-u) + uv = v(>0)$*. Note that* $0 \le u \le 1, 0 < v < \infty$*.*

$$f(u,v) = f\left(h_1^{-1}(u,v), h_2^{-1}(u,v)\right)|J|$$

$$= e^{-uv}e^{-v(1-u)}v$$

$$= ve^{-v}, \quad 0 \le u \le 1, 0 < v < \infty.$$

Functions of Several Random Variables: Method of Distribution Functions

We now discuss the distribution of Y, when Y is a function of several random variables, $Y = g(X_1, \ldots, X_n)$.

Example

Let $X_1, \ldots, X_n$ be continuous iid random variables with pdf $f(x)$ (cdf $F(x)$). Find the pdfs of

$$Y_1 = \min(X_1, \ldots, X_n) \quad \text{and} \quad Y_n = \max(X_1, \ldots, X_n).$$

Solution

For the random variable Y_1, we have

$$1 - F_{Y_1}(y) = P(Y_1 > y)$$

$$
\begin{aligned}
1 - F_{Y_1}(y) &= P(Y_1 > y) \\
&= P(X_1 > y, X_2 > y, \ldots, X_n > y) \\
&= P(X_1 > y) P(X_2 > y) \ldots P(X_n > y) \\
&\qquad \text{(because of independence)} \\
&= (1 - F(y))^n.
\end{aligned}
$$

This implies

$$F_{Y_1}(y) = 1 - (1 - F(y))^n$$

and

$$f_{Y_1}(y) = n(1 - F(y))^{n-1} f(y).$$

Consider Y_n*. Its cdf is given by*

$$F_{Y_n}(y) = P(Y_n \le y) = (F(y))^n.$$

Suppose we want the marginal $f_V(v)$ and $f_U(v)$, that is,

$$f_V(v) = \int_0^1 ve^{-v} du = ve^{-v}, \quad 0 < v < \infty$$

and

$$f_U(u) = \int_0^\infty ve^{-v} dv = 1, \quad 0 \leq u \leq 1.$$

Sometimes the expressions for two variables, U and V, may not be given. Only one expression is available. In that case, call the given expression of X and Y as U, and define $V = Y$. Then, we can use the previous method to first find the joint density and then find the marginal to obtain the pdf of U. The following example demonstrates the method.

Example :

Let X and Y be independent random variables uniformly distributed on [0, 1]. Find the distribution of $X+Y$.

Solution

Let

$$U = X + Y,$$

$$V = Y,$$

$$f(x, y) = 1, \quad 0 \le x \le 1, 0 \le y \le 1,$$

$$X = U - V,$$

$$Y = V,$$

$$J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

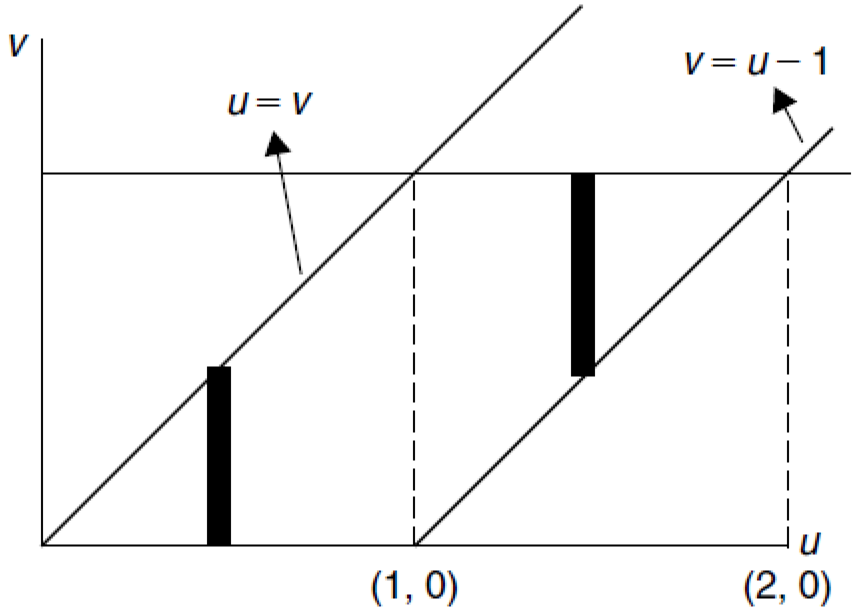
Thus, we have

$$f(u, v) = \begin{cases} 1, & 0 \le u - v \le 1, \quad 0 \le v \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

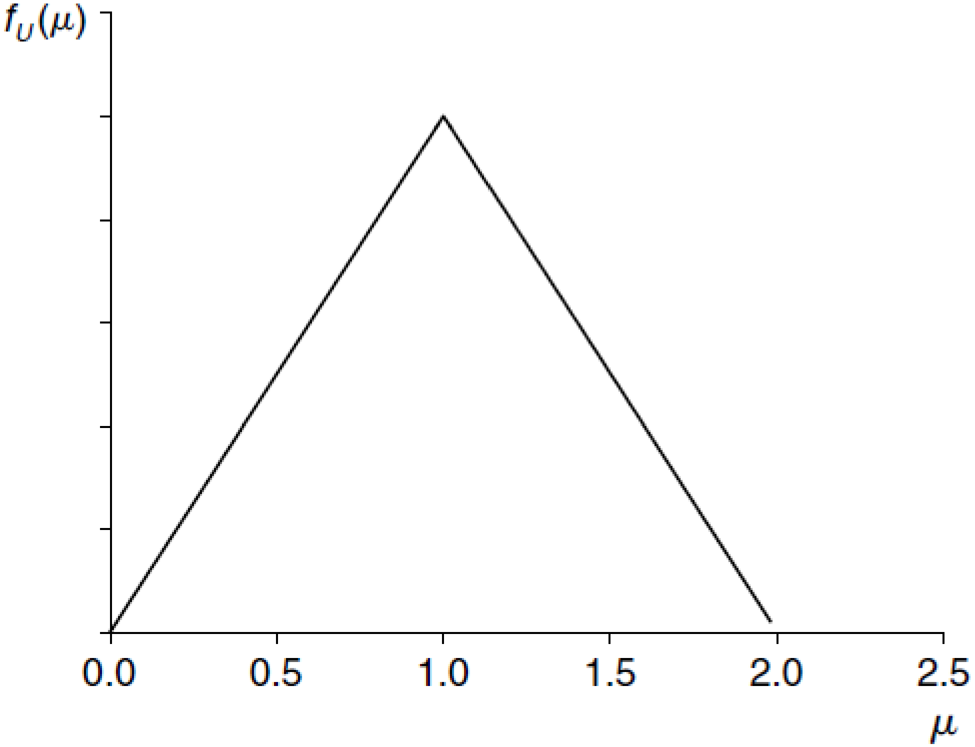
Because $\mathbf{V}$ *is the variable we introduced, to get the pdf of* $\mathbf{U}$*, we just need to find the marginal pdf from the joint pdf. From Figure 3.10, the regions of integration are* $0 \le u \le 1$*, and* $0 \le u \le 2$*. That is,*

$$f_U(u) = \int f(u, v)dv = \int 1dv$$

$$= \begin{cases} \int_0^u dv = u, & 0 \le u \le 1 \\ \int_{u-1}^1 dv = 2 - u, & 0 \le u \le 2. \end{cases}$$



■ **FIGURE** The regions of integration.



■ **FIGURE** Graph of $f_U(u)$.

EXERCISES

1. Let X be a uniformly distributed random variable over $(0, a)$. Find the pdf of $Y = cX + d$.

2. The joint pdf of (X, Y) is

$$f(x, y) = \frac{1}{\theta^2} e^{-\frac{x+y}{\theta}}, \quad x, y > 0, \;\; \theta > 0.$$

   Find the pdf of $U = X - Y$.

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور

فراس شاكر محمود

Transformation Method of one dimensional

Theorem. Let X be a continuous random variable with probability density function f(X). Let $y = T(x)$ be an increasing (or decreasing) functon .Then the density function of the random density function of the random variable $Y = T(x)$ is given by

$$g(y) = \left|\frac{dx}{dy}\right| f(W(y))$$

Where $x = W(y)$ is the inverse function of T(x).

Proof:- suppose $y = T(x)$ is an increasing function . The distribution function G(y) of Y is given by

$$G(y) = P(Y \leq y)$$

$$= P(T(x) \leq y)$$

$$= P(X \leq W(y))$$

$$= \int_{-\infty}^{W(y)} f(x)dx.$$

Then, differentiating we get the density function of Y , which is

$$g(y) = \frac{dG(y)}{dy}$$

$$= \frac{d}{dy}\left(\int_{-\infty}^{W(y)} f(x)dx\right)$$

$$= f(W(y))\frac{dW(y)}{dy}$$

$$= f\big(W(y)\big)\frac{dx}{dy} \qquad \big(since \quad x = W(y)\big).$$

On the other hand, if $y = T(x)$ is a decreasing function, then the distribution function of Y is given by

$$G(y) = P(Y \leq y)$$

$$= P(T(x) \leq y)$$

$$= P\big(X \geq W(y)\big) \qquad (\,since\; T(x)\; is\; decreasing\,)$$

$$= 1 - P(X \leq (W(y))$$

$$= 1 - \int_{-\infty}^{W(y)} f(x)dx.$$

As before, differentiating we get the density function of Y, which is

$$g(y) = \frac{dG(y)}{dy}$$

$$= \frac{d}{dy}\left(1 - \int_{-\infty}^{W(y)} f(x)dx\right)$$

$$= -f(W(y))\frac{dW(y)}{dy}$$

$$= -f(W(y))\frac{dx}{dy} \qquad \big(since \quad x = W(y)\big).$$

Hence, combining both the cases, we get

$$g(y) = \left|\frac{dx}{dy}\right| f(W(y))$$

And the proof of the theorem is now complete .

Example: $Let\ f(x) = \frac{1}{x} \qquad x \geq 1\ Find\ the\ p.d.f\ of\ Y = x$

Solution: $f(x) = \begin{cases} \frac{1}{x} & for\ \ x \geq 1 \\ 0 & o.w \end{cases}$

$g(y) = f[\omega(y)].|J|$

$y = x \Longrightarrow x = y$

$$f[\omega(y)] = \begin{cases} \frac{1}{y} & for\ y \geq 1 \\ 0 & o.w \end{cases}$$

$$|J| = \left|\frac{dx}{dy}\right| = 1$$

$$g(y) = \begin{cases} \frac{1}{y} & for\ y \geq 1 \\ 0 & o.w \end{cases}$$

Example: **If** $x \sim f(x) = 2x\ for\ 0 < x < 1$. Find the distribution of $Y = 4x^2$.

Solution:-

$$f(x) = \begin{cases} 2x & for\ 0 < x < 1 \\ 0 & o.w \end{cases}$$

$g(y) = f[\omega(y)].|J|$

$[y = 4x^2]\ \div 4$

$$\Rightarrow x^2 = \frac{y}{4}$$

$$\Rightarrow x = \frac{1}{2}\sqrt{y}$$

$$f[\omega(y)] = \begin{cases} 2\frac{\sqrt{y}}{2} & for 0 \leq y \leq 4 \\ 0 & o.w \end{cases}$$

$$\Rightarrow f[\omega(y)] = \begin{cases} \sqrt{y} & for 0 \leq y \leq 4 \\ 0 & o.w \end{cases}$$

$$|J| = \left|\frac{dx}{dy}\right| = \frac{1}{4\sqrt{y}}$$

$$g(y) = \begin{cases} \dfrac{\sqrt{y}}{4\sqrt{y}} & for\ 0 \leq y \leq 4 \\ 0 & o.w \end{cases}$$

$$g(y) = \begin{cases} \dfrac{1}{4} & for\ 0 \leq y \leq 4 \\ 0 & o.w \end{cases} \qquad g(y) \sim uniform(0,4)$$

Example: If the p.d.f. of x is $f(x) = 2xe^{-x^2}\ 0 < x < \infty$. Determine the p. d. f. of $y = x^2$.

Solution:-

$$f(x) = \begin{cases} 2xe^{-x^2} & for 0 \leq x < \infty \\ 0 & o.w \end{cases}$$

$$g(y) = f[\omega(y)].|J|$$

$$y = x^2 \quad \Rightarrow x = \sqrt{y}$$

$$f[\omega(y)] = \begin{cases} 2\sqrt{y}\, e^{-y} & for\ 0 \leq y < \infty \\ 0 & o.w \end{cases}$$

$$|J| = \left|\frac{dx}{dy}\right| = \frac{1}{2\sqrt{y}}$$

$$g(y) = \begin{cases} 2\sqrt{y}\, e^{-y} \dfrac{1}{2\sqrt{y}} & for\ 0 \leq y < \infty \\ 0 & o.w \end{cases}$$

$$g(y) = \begin{cases} e^{-y} & for\ 0 \leq y < \infty \\ 0 & o.w \end{cases} \qquad g(y) \sim Gamma(1,1)$$

Example: Let $x \sim uniform(0, \alpha)$. Determine the p. d. f. of $Y = cx + d$.

Solution:-

$$f(x) = \begin{cases} \dfrac{1}{\propto} & for 0 \leq x \leq \propto \\ 0 & o.w \end{cases}$$

$$g(y) = f[\omega(y)].|J|$$

$$y = [cx + d] \quad \div c$$

$$x = \frac{y-d}{c}$$

$$f[\omega(y)] = \begin{cases} \frac{1}{\propto} & for\ d \leq y \leq c \propto +d \\ 0 & o.w \end{cases}$$

$$|J| = \left|\frac{dx}{dy}\right| = \frac{1}{c}$$

$$g(y) = \begin{cases} \frac{1}{c \propto} & for\ d \leq y \leq c \propto +d \\ 0 & o.w \end{cases}$$

<u>Example</u>: Let $x \sim uniform\ (0,2)$. Find the p.d.f. of $Y = X^2$

<u>Solution:-</u>

$$f(x) = \begin{cases} \frac{1}{2} & for\ 0 \leq x \leq 2 \\ 0 & o.w \end{cases}$$

$$y = x^2 \Rightarrow x = \sqrt{y}$$

$$g(y) = f[\omega(y)].|J|$$

$$f[\omega(y)] = \begin{cases} \frac{1}{2} & for\ 0 \leq y \leq 4 \\ 0 & o.w \end{cases}$$

$$|J| = \left|\frac{dx}{dy}\right| = \frac{1}{2\sqrt{y}}$$

$$g(y) = \begin{cases} \frac{1}{2}\ \frac{1}{2\sqrt{y}} & for\ 0 \leq y \leq 4 \\ 0 & o.w \end{cases}$$

$$g(y) = \begin{cases} \frac{1}{4\sqrt{y}} & for\ 0 \leq y \leq 4 \\ 0 & o.w \end{cases}$$

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

Lecture Note On Mathematical Statistics 1

B.Sc. in Mathematics

Fourth Stage

Assist. Prof. Dr. Feras Shaker Mahmood

Transformation Methods of Two Dimensional

المحاضرة الرابعة

الكورس الاول

When two random variables are involved, many interesting problems can result. In the case of a single-valued inverse, the rule is about the same as that in the one-variable case, with the derivative being replaced by the Jacobian. That is, if X_1 and X_2 are two continuous-type random variables with joint pdf $f(x_1, x_2)$, and if $Y_1 = u_1(X_1, X_2)$, $Y_2 = u_2(X_1, X_2)$ has the single-valued inverse $X_1 = v_1(Y_1, Y_2)$, $X_2 = v_2(Y_1, Y_2)$, then the joint pdf of Y_1 and Y_2 is

$$g(y_1, y_2) = |J|\, f\left[v_1(y_1, y_2), v_2(y_1, y_2)\right], \qquad (y_1, y_2) \in S_Y,$$

where the Jacobian J is the determinant

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\[2ex] \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix}.$$

Of course, we find the support S_Y of Y_1, Y_2 by considering the mapping of the support S_X of X_1, X_2 under the transformation $y_1 = u_1(x_1, x_2)$, $y_2 = u_2(x_1, x_2)$. This method of finding the distribution of Y_1 and Y_2 is called the **change-of-variables technique**.

It is often the mapping of the support S_X of X_1, X_2 into that (say, S_Y) of Y_1, Y_2 which causes the biggest challenge. That is, in most cases, it is easy to solve for x_1 and x_2 in terms of y_1 and y_2, say,

$$x_1 = v_1(y_1, y_2), \qquad x_2 = v_2(y_1, y_2),$$

and then to compute the Jacobian

$$J = \begin{vmatrix} \dfrac{\partial v_1(y_1,y_2)}{\partial y_1} & \dfrac{\partial v_1(y_1,y_2)}{\partial y_2} \\ \\ \dfrac{\partial v_2(y_1,y_2)}{\partial y_1} & \dfrac{\partial v_2(y_1,y_2)}{\partial y_2} \end{vmatrix}.$$

However, the mapping of $(x_1,x_2) \in S_X$ into $(y_1,y_2) \in S_Y$ can be more difficult. Let us consider two simple examples.

Let X_1, X_2 have the joint pdf

$$f(x_1,x_2) = 2, \qquad 0 < x_1 < x_2 < 1.$$

Consider the transformation

$$Y_1 = \frac{X_1}{X_2}, \qquad Y_2 = X_2.$$

It is certainly easy enough to solve for x_1 and x_2, namely,

$$x_1 = y_1 y_2, \qquad x_2 = y_2,$$

and compute

$$J = \begin{vmatrix} y_2 & y_1 \\ \\ 0 & 1 \end{vmatrix} = y_2.$$

Let X_1 and X_2 be independent random variables, each with pdf

$$f(x) = e^{-x}, \qquad 0 < x < \infty.$$

Hence, their joint pdf is

$$f(x_1)f(x_2) = e^{-x_1-x_2}, \qquad 0 < x_1 < \infty, \ 0 < x_2 < \infty.$$

Let us consider

$$Y_1 = X_1 - X_2, \qquad Y_2 = X_1 + X_2.$$

Thus,

$$x_1 = \frac{y_1 + y_2}{2}, \qquad x_2 = \frac{y_2 - y_1}{2},$$

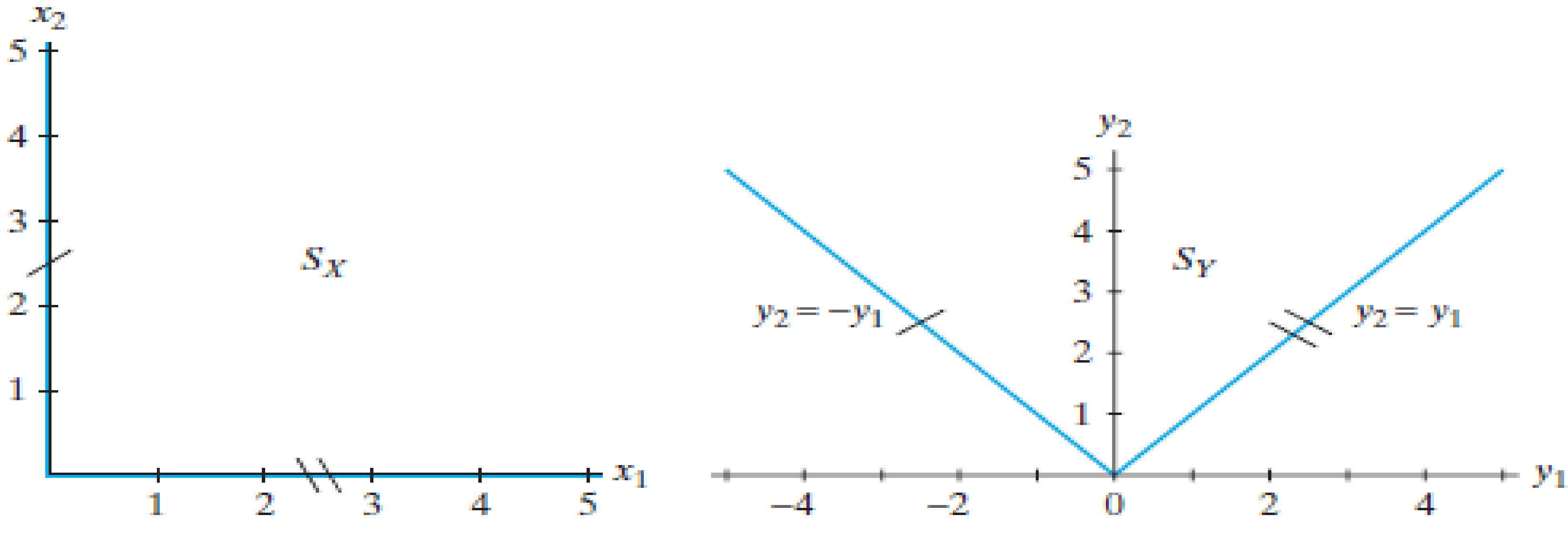
with

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{2}.$$

The region S_X is depicted . The line segments on the boundary, namely, $x_1 = 0, \ 0 < x_2 < \infty$, and $x_2 = 0, \ 0 < x_1 < \infty$, map into the line segments $y_1 + y_2 = 0, \ y_2 > y_1$ and $y_1 = y_2, \ y_2 > -y_1$, respectively. These are shown in Figure 5.2-2(b), and the support of S_Y is depicted there. Since the region S_Y is not bounded by horizontal and vertical line segments, Y_1 and Y_2 are dependent.

The joint pdf of Y_1 and Y_2 is

$$g(y_1, y_2) = \frac{1}{2} e^{-y_2}, \qquad -y_2 < y_1 < y_2, \ 0 < y_2 < \infty.$$



(a) Support of X_1, X_2 (b) Support of Y_1, Y_2

Figure Mapping from x_1, x_2 to y_1, y_2

The probability $P(Y_1 \geq 0, Y_2 \leq 4)$ is given by

$$\int_0^4 \int_{y_1}^4 \frac{1}{2} e^{-y_2}\, dy_2\, dy_1 \qquad \text{or} \qquad \int_0^4 \int_0^{y_2} \frac{1}{2} e^{-y_2}\, dy_1\, dy_2.$$

While neither of these integrals is difficult to evaluate, we choose the latter one to obtain

$$\int_0^4 \frac{1}{2} y_2 e^{-y_2}\, dy_2 = \left[\frac{1}{2}(-y_2)e^{-y_2} - \frac{1}{2} e^{-y_2}\right]_0^4$$
$$= \frac{1}{2} - 2e^{-4} - \frac{1}{2} e^{-4} = \frac{1}{2}\left[1 - 5e^{-4}\right].$$

The marginal pdf of Y_2 is

$$g_2(y_2) = \int_{-y_2}^{y_2} \frac{1}{2} e^{-y_2}\, dy_1 = y_2 e^{-y_2}, \qquad 0 < y_2 < \infty.$$

This is a gamma pdf with shape parameter 2 and scale parameter 1. The pdf of Y_1 is

$$g_1(y_1) = \begin{cases} \displaystyle\int_{-y_1}^{\infty} \frac{1}{2} e^{-y_2}\, dy_2 = \frac{1}{2} e^{y_1}, & -\infty < y_1 \le 0, \\[2ex] \displaystyle\int_{y_1}^{\infty} \frac{1}{2} e^{-y_2}\, dy_2 = \frac{1}{2} e^{-y_1}, & 0 < y_1 < \infty. \end{cases}$$

That is, the expression for $g_1(y_1)$ depends on the location of y_1, although this could be written as

$$g_1(y_1) = \frac{1}{2} e^{-|y_1|}, \qquad -\infty < y_1 < \infty,$$

which is called a **double exponential** pdf, or sometimes the **Laplace** pdf. ■

Example Let X_1 and X_2 have independent gamma distributions with parameters α, θ and β, θ, respectively. That is, the joint pdf of X_1 and X_2 is

$$f(x_1, x_2) = \frac{1}{\Gamma(\alpha)\Gamma(\beta)\theta^{\alpha+\beta}} x_1^{\alpha-1} x_2^{\beta-1} \exp\left(-\frac{x_1 + x_2}{\theta}\right), \; 0 < x_1 < \infty, \; 0 < x_2 < \infty.$$

Consider

$$Y_1 = \frac{X_1}{X_1 + X_2}, \qquad Y_2 = X_1 + X_2,$$

or, equivalently,

$$X_1 = Y_1 Y_2, \qquad X_2 = Y_2 - Y_1 Y_2.$$

The Jacobian is

$$J = \begin{vmatrix} y_2 & y_1 \\ -y_2 & 1 - y_1 \end{vmatrix} = y_2(1 - y_1) + y_1 y_2 = y_2.$$

Thus, the joint pdf $g(y_1, y_2)$ of Y_1 and Y_2 is

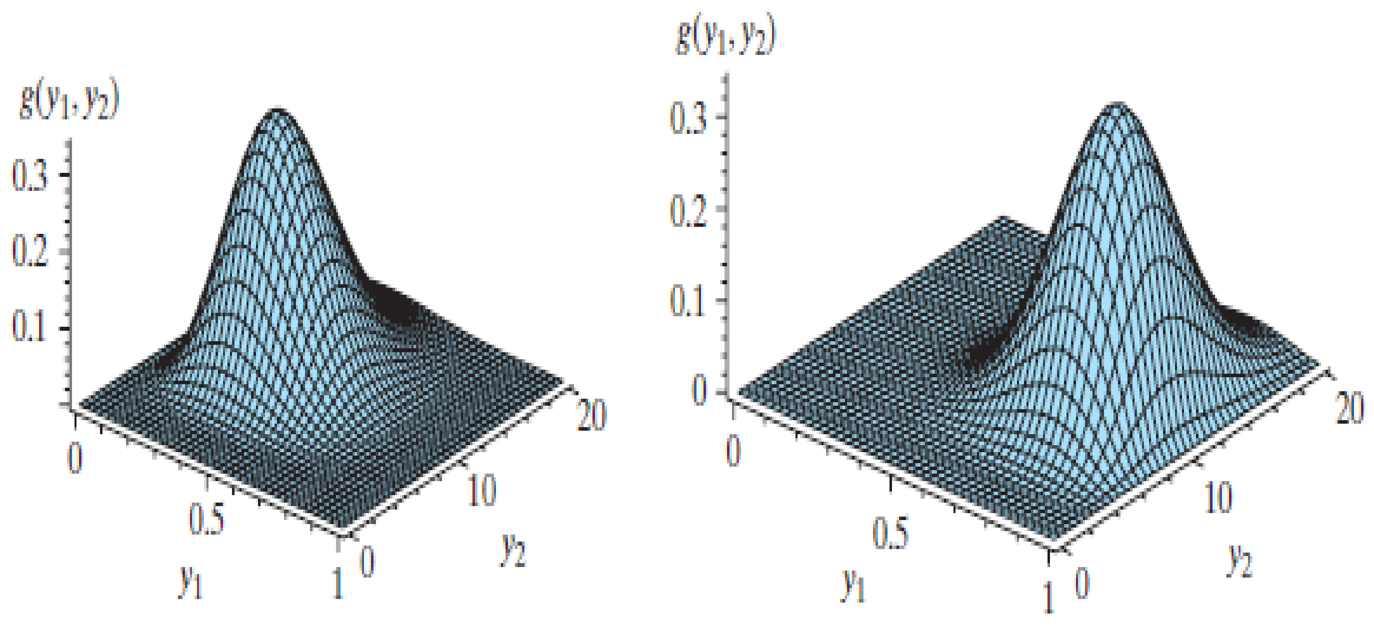
$$g(y_1, y_2) = |y_2| \frac{1}{\Gamma(\alpha)\Gamma(\beta)\theta^{\alpha+\beta}} (y_1 y_2)^{\alpha-1} (y_2 - y_1 y_2)^{\beta-1} e^{-y_2/\theta},$$

where the support is $0 < y_1 < 1$, $0 < y_2 < \infty$, which is the mapping of $0 < x_i < \infty$, $i = 1, 2$. To see the shape of this joint pdf, $z = g(y_1, y_2)$ is graphed in Figure 5.2-3(a) with $\alpha = 4$, $\beta = 7$, and $\theta = 1$ and in Figure (b) with $\alpha = 8$, $\beta = 3$, and $\theta = 1$. To find the marginal pdf of Y_1, we integrate this joint pdf on y_2. We see that the marginal pdf of Y_1 is

$$g_1(y_1) = \frac{y_1^{\alpha-1}(1 - y_1)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \frac{y_2^{\alpha+\beta-1}}{\theta^{\alpha+\beta}} e^{-y_2/\theta} \, dy_2.$$

But the integral in this expression is that of a gamma pdf with parameters $\alpha + \beta$ and θ, except for $\Gamma(\alpha + \beta)$ in the denominator; hence, the integral equals $\Gamma(\alpha + \beta)$, and we have

$$g_1(y_1) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y_1^{\alpha-1}(1 - y_1)^{\beta-1}, \qquad 0 < y_1 < 1.$$



(a) $\alpha = 4, \beta = 7, \theta = 1$

(b) $\alpha = 8, \beta = 3, \theta = 1$

Joint pdf of $z = g(y_1, y_2)$

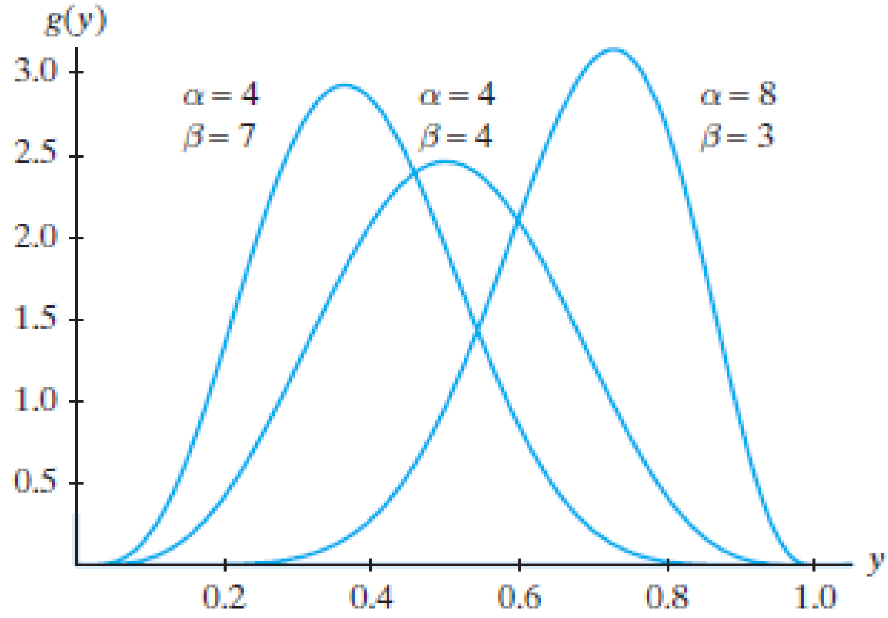


Figure Beta distribution pdfs

We say that Y_1 has a **beta pdf** with parameters α and β.

The next example illustrates the distribution function technique. You will calculate the same results in Exercise 5.2-2, but using the change-of-variable technique.

Example We let

$$F = \frac{U/r_1}{V/r_2},$$

where U and V are independent chi-square variables with r_1 and r_2 degrees of freedom, respectively. Thus, the joint pdf of U and V is

$$g(u, v) = \frac{u^{r_1/2-1} e^{-u/2}}{\Gamma(r_1/2) 2^{r_1/2}} \frac{v^{r_2/2-1} e^{-v/2}}{\Gamma(r_2/2) 2^{r_2/2}}, \qquad 0 < u < \infty,\ 0 < v < \infty.$$

In this derivation, we let $W = F$ to avoid using f as a symbol for a variable. The cdf $F(w) = P(W \le w)$ of W is

$$F(w) = P\left(\frac{U/r_1}{V/r_2} \le w\right) = P\left(U \le \frac{r_1}{r_2} w V\right)$$
$$= \int_0^\infty \int_0^{(r_1/r_2)wv} g(u, v)\, du\, dv.$$

That is,

$$F(w) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)} \int_0^\infty \left[\int_0^{(r_1/r_2)wv} \frac{u^{r_1/2-1} e^{-u/2}}{2^{(r_1+r_2)/2}}\, du \right] v^{r_2/2-1} e^{-v/2}\, dv.$$

The pdf of W is the derivative of the cdf; so, applying the fundamental theorem of calculus to the inner integral, exchanging the operations of integration and differentiation (which is permissible in this case), we have

$$f(w) = F'(w)$$

$$= \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)} \int_0^\infty \frac{[(r_1/r_2)vw]^{r_1/2-1}}{2^{(r_1+r_2)/2}} e^{-(r_1/2r_2)(vw)} \left(\frac{r_1}{r_2} v\right) v^{r_2/2-1} e^{-v/2}\, dv$$

$$= \frac{(r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)} \int_0^\infty \frac{v^{(r_1+r_2)/2-1}}{2^{(r_1+r_2)/2}} e^{-(v/2)[1+(r_1/r_2)w]}\, dv.$$

In the integral, we make the change of variable

$$y = \left(1 + \frac{r_1}{r_2} w\right) v, \qquad \text{so that} \qquad \frac{dv}{dy} = \frac{1}{1+(r_1/r_2)w}.$$

Thus, we have

$$f(w) = \frac{(r_1/r_2)^{r_1/2}\Gamma[(r_1+r_2)/2] w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)[1+(r_1 w/r_2)]^{(r_1+r_2)/2}} \int_0^\infty \frac{y^{(r_1+r_2)/2-1} e^{-y/2}}{\Gamma[(r_1+r_2)/2] 2^{(r_1+r_2)/2}}\, dy$$

$$= \frac{(r_1/r_2)^{r_1/2}\Gamma[(r_1+r_2)/2] w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)[1+(r_1 w/r_2)]^{(r_1+r_2)/2}},$$

the pdf of the $W = F$ **distribution** with r_1 and r_2 degrees of freedom. Note that the integral in this last expression for $f(w)$ is equal to 1 because the integrand is like a pdf of a chi-square distribution with $r_1 + r_2$ degrees of freedom. Graphs of pdfs for the F distribution

If all n of the distributions are the same, then the collection of n independent and identically distributed random variables, $X_1, X_2, \ldots, X_n$, is said to be a **random sample of size n from that common distribution.** If $f(x)$ is the common pmf or pdf of these n random variables, then the joint pmf or pdf is $f(x_1)f(x_2)\cdots f(x_n)$.

Example Let X_1, X_2, X_3 be a random sample from a distribution with pdf

$$f(x) = e^{-x}, \qquad 0 < x < \infty.$$

The joint pdf of these three random variables is

$$f(x_1, x_2, x_3) = (e^{-x_1})(e^{-x_2})(e^{-x_3}) = e^{-x_1-x_2-x_3}, \qquad 0 < x_i < \infty, \; i = 1, 2, 3.$$

The probability

$$P(0 < X_1 < 1, 2 < X_2 < 4, 3 < X_3 < 7)$$

$$= \left(\int_0^1 e^{-x_1}\,dx_1\right)\left(\int_2^4 e^{-x_2}\,dx_2\right)\left(\int_3^7 e^{-x_3}\,dx_3\right)$$

$$= (1 - e^{-1})(e^{-2} - e^{-4})(e^{-3} - e^{-7}),$$

because of the independence of X_1, X_2, X_3. ■

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

Lecture Note On Mathematical Statistics 1

B.Sc. in Mathematics

Fourth Stage

Assist. Prof. Dr. Feras Shaker Mahmood

Moment Generating Technique

المحاضرة الخامسة

الكورس الاول

The first three sections of this chapter presented several techniques for determining the distribution of a function of random variables with known distributions. Another technique for this purpose is the moment-generating function technique. If $Y = u(X_1, X_2, \ldots, X_n)$, we have noted that we can find $E(Y)$ by evaluating $E[u(X_1, X_2, \ldots, X_n)]$. It is also true that we can find $E[e^{tY}]$ by evaluating $E[e^{tu(X_1, X_2, \ldots, X_n)}]$. We begin with a simple example.

Example 1

Let X_1 and X_2 be independent random variables with uniform distributions on $\{1, 2, 3, 4\}$. Let $Y = X_1 + X_2$. For example, Y could equal the sum when two fair four-sided dice are rolled. The mgf of Y is

$$M_Y(t) = E\left(e^{tY}\right) = E\left[e^{t(X_1+X_2)}\right] = E\left(e^{tX_1}e^{tX_2}\right).$$

The independence of X_1 and X_2 implies that

$$M_Y(t) = E\left(e^{tX_1}\right) E\left(e^{tX_2}\right).$$

In this example, X_1 and X_2 have the same pmf, namely,

$$f(x) = \frac{1}{4}, \qquad x = 1, 2, 3, 4,$$

and thus the same mgf,

$$M_X(t) = \frac{1}{4}e^t + \frac{1}{4}e^{2t} + \frac{1}{4}e^{3t} + \frac{1}{4}e^{4t}.$$

It then follows that $M_Y(t) = [M_X(t)]^2$ equals

$$\frac{1}{16}e^{2t} + \frac{2}{16}e^{3t} + \frac{3}{16}e^{4t} + \frac{4}{16}e^{5t} + \frac{3}{16}e^{6t} + \frac{2}{16}e^{7t} + \frac{1}{16}e^{8t}.$$

Note that the coefficient of e^{bt} is equal to the probability $P(Y = b)$; for example, $4/16 = P(Y = 5)$. Thus, we can find the distribution of Y by determining its mgf. ■

Theorem 1 If $X_1, X_2, \ldots, X_n$ are independent random variables with respective moment-generating functions $M_{X_i}(t)$, $i = 1, 2, 3, \ldots, n$, where $-h_i < t < h_i, i = 1, 2, \ldots, n$, for positive numbers $h_i, i = 1, 2, \ldots, n$, then the moment-generating function of $Y = \sum_{i=1}^{n} a_i X_i$ is

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t), \quad \text{where} \quad -h_i < a_i t < h_i, \; i = 1, 2, \ldots, n.$$

Proof From Theorem 5.3-1, the mgf of Y is given by

$$\begin{aligned} M_Y(t) = E\left[e^{tY}\right] &= E\left[e^{t(a_1X_1 + a_2X_2 + \cdots + a_nX_n)}\right] \\ &= E\left[e^{a_1tX_1} e^{a_2tX_2} \cdots e^{a_ntX_n}\right] \\ &= E\left[e^{a_1tX_1}\right] E\left[e^{a_2tX_2}\right] \cdots E\left[e^{a_ntX_n}\right]. \end{aligned}$$

However, since

$$E\left(e^{tX_i}\right) = M_{X_i}(t),$$

it follows that

$$E\left(e^{a_itX_i}\right) = M_{X_i}(a_it).$$

Thus, we have

$$M_Y(t) = M_{X_1}(a_1t) M_{X_2}(a_2t) \cdots M_{X_n}(a_nt) = \prod_{i=1}^{n} M_{X_i}(a_it). \qquad \square$$

Corollary 1

If $X_1, X_2, \ldots, X_n$ are observations of a random sample from a distribution with moment-generating function $M(t)$, where $-h < t < h$, then

(a) the moment-generating function of $Y = \sum_{i=1}^{n} X_i$ is

$$M_Y(t) = \prod_{i=1}^{n} M(t) = [M(t)]^n, \quad -h < t < h;$$

(b) the moment-generating function of $\overline{X} = \sum_{i=1}^{n} (1/n) X_i$ is

$$M_{\overline{X}}(t) = \prod_{i=1}^{n} M\left(\frac{t}{n}\right) = \left[M\left(\frac{t}{n}\right)\right]^n, \quad -h < \frac{t}{n} < h.$$

Proof For (a), let $a_i = 1$, $i = 1, 2, \ldots, n$, in Theorem 5.4-1. For (b), take $a_i = 1/n$, $i = 1, 2, \ldots, n$. ◀

The next two examples and the exercises give some important applications of Theorem 1 and its corollary. Recall that the mgf, once found, uniquely determines the distribution of the random variable under consideration.

Example 2

Let $X_1, X_2, \ldots, X_n$ denote the outcomes of n Bernoulli trials, each with probability of success p. The mgf of $X_i, i = 1, 2, \ldots, n$, is

$$M(t) = q + pe^t, \quad -\infty < t < \infty.$$

If

$$Y = \sum_{i=1}^{n} X_i,$$

then

$$M_Y(t) = \prod_{i=1}^{n} (q + pe^t) = (q + pe^t)^n, \qquad -\infty < t < \infty.$$

Thus, we again see that Y is $b(n, p)$.

Example 3

Let X_1, X_2, X_3 be the observations of a random sample of size $n = 3$ from the exponential distribution having mean θ and, of course, mgf $M(t) = 1/(1 - \theta t)$, $t < 1/\theta$. The mgf of $Y = X_1 + X_2 + X_3$ is

$$M_Y(t) = \left[(1 - \theta t)^{-1}\right]^3 = (1 - \theta t)^{-3}, \qquad t < 1/\theta,$$

which is that of a gamma distribution with parameters $\alpha = 3$ and θ. Thus, Y has this distribution. On the other hand, the mgf of $\overline{X}$ is

$$M_{\overline{X}}(t) = \left[\left(1 - \frac{\theta t}{3}\right)^{-1}\right]^3 = \left(1 - \frac{\theta t}{3}\right)^{-3}, \qquad t < 3/\theta.$$

Hence, the distribution of $\overline{X}$ is gamma with the parameters $\alpha = 3$ and $\theta/3$, respectively.

Theorem 2 Let $X_1, X_2, \ldots, X_n$ be independent chi-square random variables with $r_1, r_2, \ldots, r_n$ degrees of freedom, respectively. Then $Y = X_1+X_2+\cdots+X_n$ is $\chi^2(r_1+r_2+\cdots+r_n)$.

Proof By Theorem 5.4-1 with each $a = 1$, the mgf of Y is

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = (1-2t)^{-r_1/2}(1-2t)^{-r_2/2}\cdots(1-2t)^{-r_n/2}$$
$$= (1-2t)^{-\Sigma r_i/2}, \qquad \text{with } t < 1/2,$$

which is the mgf of a $\chi^2(r_1+r_2+\cdots+r_n)$. Thus, Y is $\chi^2(r_1+r_2+\cdots+r_n)$. □

The next two corollaries combine and extend the results of Theorems 1 and 2 and give one interpretation of degrees of freedom.

Corollary 2 Let $Z_1, Z_2, \ldots, Z_n$ have standard normal distributions, $N(0,1)$. If these random variables are independent, then $W = Z_1^2 + Z_2^2 + \cdots + Z_n^2$ has a distribution that is $\chi^2(n)$.

Proof By Theorem 1, Z_i^2 is $\chi^2(1)$ for $i = 1, 2, \ldots, n$. From Theorem 2, with $Y = W$ and $r_i = 1$, it follows that W is $\chi^2(n)$. ◀

Corollary 3

If $X_1, X_2, \ldots, X_n$ are independent and have normal distributions $N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots, n$, respectively, then the distribution of

$$W = \sum_{i=1}^{n} \frac{(X_i - \mu_i)^2}{\sigma_i^2}$$

is $\chi^2(n)$.

Proof This follows from Corollary 2 since $Z_i = (X_i - \mu_i)/\sigma_i$ is $N(0,1)$, and thus

$$Z_i^2 = \frac{(X_i - \mu_i)^2}{\sigma_i^2}$$

is $\chi^2(1)$, $i = 1, 2, \ldots, n$. ◀

Theorem 3

If $X_1, X_2, \ldots, X_n$ are n mutually independent normal variables with means $\mu_1, \mu_2, \ldots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, respectively, then the linear function

$$Y = \sum_{i=1}^{n} c_i X_i$$

has the normal distribution

$$N\left(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2\right).$$

Proof By Theorem 5.4-1, we have, with $-\infty < c_i t < \infty$, or $-\infty < t < \infty$,

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(c_i t) = \prod_{i=1}^{n} \exp\left(\mu_i c_i t + \sigma_i^2 c_i^2 t^2/2\right)$$

because $M_{X_i}(t) = \exp(\mu_i t + \sigma_i^2 t^2/2)$, $i = 1, 2, \ldots, n$. Thus,

$$M_Y(t) = \exp\left[\left(\sum_{i=1}^{n} c_i \mu_i\right) t + \left(\sum_{i=1}^{n} c_i^2 \sigma_i^2\right)\left(\frac{t^2}{2}\right)\right].$$

This is the mgf of a distribution that is

$$N\left(\sum_{i=1}^{n} c_i\mu_i, \sum_{i=1}^{n} c_i^2\sigma_i^2\right).$$

Thus, Y has this normal distribution. □

From Theorem 3, we observe that the difference of two independent normally distributed random variables, say, $Y = X_1 - X_2$, has the normal distribution $N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

Example 4

Let X_1 and X_2 equal the number of pounds of butterfat produced by two Holstein cows (one selected at random from those on the Koopman farm and one selected at random from those on the Vliestra farm, respectively) during the 305-day lactation period following the births of calves. Assume that the distribution of X_1 is $N(693.2, 22820)$ and the distribution of X_2 is $N(631.7, 19205)$. Moreover, let X_1 and X_2 be independent. We shall find $P(X_1 > X_2)$. That is, we shall find the probability that the butterfat produced by the Koopman farm cow exceeds that produced by the Vliestra farm cow. (Sketch pdfs on the same graph for these two normal distributions.) If we let $Y = X_1 - X_2$, then the distribution of Y is $N(693.2 - 631.7, 22820 + 19205)$. Thus,

$$P(X_1 > X_2) = P(Y > 0) = P\left(\frac{Y - 61.5}{\sqrt{42025}} > \frac{0 - 61.5}{205}\right)$$

$$= P(Z > -0.30) = 0.6179.$$

Theorem 4

Let $X_1, X_2, \ldots, X_n$ be observations of a random sample of size n from the normal distribution $N(\mu, \sigma^2)$. Then the sample mean,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

and the sample variance,

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

are independent and

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} \quad \text{is} \quad \chi^2(n-1).$$

Proof We are not prepared to prove the independence of $\overline{X}$ and S^2 at this time , so we accept it without proof here. To prove the second part, note that

$$W = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2 = \sum_{i=1}^{n}\left[\frac{(X_i - \overline{X}) + (\overline{X} - \mu)}{\sigma}\right]^2$$

$$= \sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 + \frac{n(\overline{X} - \mu)^2}{\sigma^2} \qquad (5.5\text{-}1)$$

because the cross-product term is equal to

$$2\sum_{i=1}^{n}\frac{\left(\overline{X}-\mu\right)\left(X_i-\overline{X}\right)}{\sigma^2}=\frac{2\left(\overline{X}-\mu\right)}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)=0.$$

But $Y_i = (X_i - \mu)/\sigma$, $i = 1, 2, \ldots, n$, are standardized normal variables that are independent. Hence, $W = \sum_{i=1}^{n} Y_i^2$ is $\chi^2(n)$ by Corollary 5.4-3. Moreover, since $\overline{X}$ is $N(\mu, \sigma^2/n)$, it follows that

$$Z^2=\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2=\frac{n(\overline{X}-\mu)^2}{\sigma^2}$$

is $\chi^2(1)$ by Theorem 2. In this notation,

$$W=\frac{(n-1)S^2}{\sigma^2}+Z^2.$$

However, from the fact that $\overline{X}$ and S^2 are independent, it follows that Z^2 and S^2 are also independent. In the mgf of W, this independence permits us to write

$$E\left[e^{tW}\right]=E\left[e^{t\{(n-1)S^2/\sigma^2+Z^2\}}\right]=E\left[e^{t(n-1)S^2/\sigma^2}e^{tZ^2}\right]$$
$$=E\left[e^{t(n-1)S^2/\sigma^2}\right]E\left[e^{tZ^2}\right].$$

Since W and Z^2 have chi-square distributions, we can substitute their mgfs to obtain

$$(1-2t)^{-n/2}=E\left[e^{t(n-1)S^2/\sigma^2}\right](1-2t)^{-1/2}.$$

Equivalently, we have

$$E\left[e^{t(n-1)S^2/\sigma^2}\right]=(1-2t)^{-(n-1)/2},\qquad t<\frac{1}{2}.$$

This, of course, is the mgf of a $\chi^2(n-1)$-variable; accordingly, $(n-1)S^2/\sigma^2$ has that distribution. □

Theorem 5 **(Student's t distribution)** Let

$$T = \frac{Z}{\sqrt{U/r}},$$

where Z is a random variable that is $N(0,1)$, U is a random variable that is $\chi^2(r)$, and Z and U are independent. Then T has a t distribution with pdf

$$f(t) = \frac{\Gamma((r+1)/2)}{\sqrt{\pi r}\,\Gamma(r/2)} \frac{1}{(1+t^2/r)^{(r+1)/2}}, \qquad -\infty < t < \infty.$$

Proof The joint pdf of Z and U is

$$g(z,u) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \frac{1}{\Gamma(r/2)2^{r/2}} u^{r/2-1} e^{-u/2}, \qquad -\infty < z < \infty,\ 0 < u < \infty.$$

The cdf $F(t) = P(T \le t)$ of T is given by

$$\begin{aligned} F(t) &= P\left(Z/\sqrt{U/r} \le t\right) \\ &= P\left(Z \le \sqrt{U/r}\,t\right) \\ &= \int_0^\infty \int_{-\infty}^{\sqrt{(u/r)}\,t} g(z,u)\,dz\,du. \end{aligned}$$

That is,

$$F(t) = \frac{1}{\sqrt{\pi}\,\Gamma(r/2)} \int_0^\infty \left[\int_{-\infty}^{\sqrt{(u/r)}\,t} \frac{e^{-z^2/2}}{2^{(r+1)/2}}\,dz\right] u^{r/2-1} e^{-u/2}\,du.$$

The pdf of T is the derivative of the cdf; so, applying the fundamental theorem of calculus to the inner integral (interchanging the derivative and integral operators is permitted here), we find that

$$f(t) = F'(t) = \frac{1}{\sqrt{\pi}\,\Gamma(r/2)} \int_0^\infty \frac{e^{-(u/2)(t^2/r)}}{2^{(r+1)/2}} \sqrt{\frac{u}{r}}\, u^{r/2-1} e^{-u/2}\, du$$

$$= \frac{1}{\sqrt{\pi r}\,\Gamma(r/2)} \int_0^\infty \frac{u^{(r+1)/2-1}}{2^{(r+1)/2}} e^{-(u/2)(1+t^2/r)}\, du.$$

In the integral, make the change of variables

$$y = (1 + t^2/r)u, \qquad \text{so that} \qquad \frac{du}{dy} = \frac{1}{1 + t^2/r}.$$

Thus,

$$f(t) = \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)} \left[\frac{1}{(1 + t^2/r)^{(r+1)/2}}\right] \int_0^\infty \frac{y^{(r+1)/2-1}}{\Gamma[(r+1)/2]\, 2^{(r+1)/2}} e^{-y/2}\, dy.$$

The integral in this last expression for $f(t)$ is equal to 1 because the integrand is like the pdf of a chi-square distribution with $r + 1$ degrees of freedom. Hence, the pdf is

$$f(t) = \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)} \frac{1}{(1 + t^2/r)^{(r+1)/2}}, \qquad -\infty < t < \infty. \qquad \square$$

Example 5

Let the distribution of T be $t(11)$. Then

$$t_{0.05}(11) = 1.796 \qquad \text{and} \qquad -t_{0.05}(11) = -1.796.$$

Thus,

$$P(-1.796 \le T \le 1.796) = 0.90.$$

We can also find values of the cdf such as

$$P(T \le 2.201) = 0.975 \qquad \text{and} \qquad P(T \le -1.363) = 0.10.$$

We can use the results of Corollary 3 and Theorems 3 and 5 to construct an important T random variable. Given a random sample $X_1, X_2, \ldots, X_n$ from a normal distribution, $N(\mu, \sigma^2)$, let

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \qquad \text{and} \qquad U = \frac{(n-1)S^2}{\sigma^2}.$$

Then the distribution of Z is $N(0,1)$ by Corollary 3. Theorem 3 tells us that the distribution of U is $\chi^2(n-1)$ and that Z and U are independent. Thus,

$$T = \frac{\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/(n-1)}} = \frac{\overline{X} - \mu}{S/\sqrt{n}} \tag{5.5-2}$$

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور

فراس شاكر محمود

Moment Generating Function Method :-

Theorem :

Let m_x (t) and m_y (t) denote the moment - generating functions of random variables X and Y , respectively . If both moment - generating functions exist and m_x (t) $= m_y$ (t) for all values of t , then X and Y have the same probability distribution.

Example: Let X and Y be independent random variables Gamma distributed on [a , 1] . Find the distribution of Z=X + Y .

Solution :

$M_x (t) = (1-t)^{-a}$, $M_y (t) = (1-t)^{-a}$

$M_{x+y} (t) = E(e^{tx+ty}) = E(e^{(tx)}. \ e^{(ty)}) = E(e^{(tx)}). E(e^{(ty)}) = M_x (t) M_y (t)$

$=(1-t)^{-a} . (1-t)^{-a}$

$= \ (1-t)^{-2a}$

That is a moment generating of Z is Gamma (2a , 1) , Thus

Z= X + Y ~ Gamma (2a, 1) .

In general if a identically independent r .v's x_i ~ Gamma (a, β) , ∀i = 1,2 ….. n. Find the p. d. f of Y = $\sum_{i=0}^{n} x_i$

Mx_i (t) $=$ (1 $-$ Bt) $^{-ai}$, ∀i = 1,2 , ... n

$M_y (t) = E(e^{ty}) = E(e^{t(x1+x2+\cdots+xn)} = E(e^{(tx1+tx2+\cdots+txn)})=E(e^{(tx1)} . e^{(tx2)} \ .... e^{(txn)})$

$= E(e^{(tx1)}) E(e^{(tx2)}) ... ... E(e^{(txn)}) = M_{x1} (t). M_{x2} (t) ..... M_{xn} (t)$

$= (1-\beta t)^{-a1} . (1-\beta t)^{-a2} \dots (1-\beta t)^{-an}$

$= (1-\beta t)^{-\sum_{i=1}^{n} ai}$

That is a moment generating of Gamma ($\sum_{i=1}^{n} ai$, β)

Y ~ Gamma ($\sum_{i=1}^{n} ai$, β)

Example :-

Let X_1 , X_2 , ... , X_n , be independent and identically distributed random variables such that Xi ~ N ($\mu i . \sigma i^2$) , ∀i = 1,2 , ……... , n , Find the p.d.f of Y=$\sum_{i=1}^{n} ai\ Xi\ \ .a\ is$ constant .

M_{xi} (t) = EXP $\left\{\mu i\ t + \frac{1}{2}\ \sigma i^2 t^2\right\}$; i = 1 , 2 ……… n

M_y (t) = = E(e^{ty}) = E($e^{t(a1x1+a2x2+\cdots an\ xn)}$ = E($e^{(ta1x1+ta2x2+\cdots tan\ xn)}$)

= E($e^{(ta_1x_1)}$) E($e^{(ta_2x_2)}$) ... E($e^{(ta_nx_n)}$)= M_{x1} (a_1t) M_{x2} (a_2t) …M_{xn}(a_nt)

= EXP $\left\{\mu 1\ a1\ t + \frac{1}{2}\ a1^2\sigma 1^2 t^2\right\}$ EXP $\left\{\mu 2\ \ a2\ t + \frac{1}{2}\ \sigma 2^2 a2^2 t^2\right\}$...EXP $\left\{\mu n\ an\ t + \frac{1}{2}\ \sigma n^2 an^2 t^2\right\}$

That is a moment generating function of Y~ $N\left[\!\left[\sum_{i=1}^{n} ai\ \mu i\ \ ,\ \ \sum_{i=1}^{n} a_i^2\sigma_i^2\right]\!\right]$

For a special case that is if X , Y ~ $N\ (\ \mu\ , \sigma^2\ )$? then X - Y ~ N ($\mu -\ \mu\ , \sigma^2 +\ \sigma^2$)

i. e X - Y ~ N (0,2σ^2)

Example:

Let Y1 , Y2 …… Yn be independent and identically distributed random variables such that for $0 < P\ < 1.\ P\ (Yi = 1\ ) = p$ and $p(\ Yi = 0\ ) = q = 1 - p$ such random variables are called random variables. W = Y1 +Y2 ….. +Yn. What is the distribution of W?

Solution :

My (t) = (pe^t + q)

Mw (t) = E (e^{tw}) = E ($e^{t(\ y1+y2+\ \ldots\ldots+yn\ )}$) = E ($e^{(\ ty1+ty2+\ \ldots\ldots+tyn\ )}$)

= E (e^{ty1}) E (e^{ty2}) ….E (e^{tyn}) = My1 (t) My2 (t) …. Myn (t)

=(pe^t + q).(pe^t + q) …..(pe^t + q) = (pe^t + q$)^n$

That is a moment generating of b(n.p) .Thus W ~ b(n .p).

Example:

If X ~ N (0, 1) thenY= X^2 ~χ_1^2?

Solution : Let Y = X^2 , f (x) =$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, - ∞ $< x < \infty$

M_y (t) = E (e^{ty}) = E (e^{tx^2}) = $\int_{-\infty}^{\infty} e^{tx^2}\ \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\ dx\ ) = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}+tx^2}\ dx$

$= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}(1-2t)}\ dx = \frac{1}{(1-2t)^{\frac{1}{2}}}\ \int_{-\infty}^{\infty}\frac{(\ 1-2t)^{\frac{1}{2}}}{\sqrt{2\pi}}\ \ e^{\frac{x^2(1-2t)}{2}}\ \ dx = (\ 1 - 2t)^{\frac{-1}{2}}$

That is a moment generating of χ_1^2. Thus $Y\sim \chi_1^2$.

In general If $x_1, x_2, \dots, x_n \sim N(0,1)$, then $Y = x_1^2 + x_2^2 + \dots + x_n^2 \sim \chi_n^2$.

Example: If x~ N(μ , σ^2) , then Y = $\frac{X-\mu}{\sigma}$ ~ N (0 , 1)

Solution: $x \sim N(\mu, \sigma^2) \quad \rightarrow f(x) = \frac{1}{\sqrt{2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma}}$ $, -\infty < x < \infty$

$Y = \frac{x-\mu}{\sigma} \rightarrow \sigma y = x - \mu \rightarrow \sigma dy = dx$

$$My(t) = e^{ty} = \int_{-\infty}^{\infty} e^{t\left(\frac{x-\mu}{\sigma}\right)} f(x)dx = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{t\left(\frac{x-\mu}{\sigma}\right) - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} e^{ty} . e^{-\frac{y2}{2}} . \sigma dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(y^2+2ty)}{2}} dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(\frac{(y^2+2ty+t^2-t^2)}{2}} dy = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(\frac{(y-t)^2}{2}} dy$$

Let y –t = w → $dy = dw$, My (t) = $e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw = e^{\frac{t^2}{2}}$

Y ~ N (0 , 1)

The Distribution of $\bar{x}$: Let x_1 ,x_2.....$and\ x_n$ are independent and identically distributed normal random variables with mean μ and variance σ^2 , then the way to find the Distribution of $\bar{X}$ is

m_x (t) =EXP $\left\{\mu t + \frac{1}{2}\sigma^2 r^2\right\}$

$$x_{\bar{x}}(t) = E\left(e^{t\bar{x}}\right) = E\left(e^{\frac{t}{n}(x_{1}+x_{2}+\dots+Xn}\right) = E\left(e^{\frac{t}{n}}x_1 + \frac{t}{n}x_2 + \dots + \frac{t}{n}x_n\right)$$

$$= E\left(e^{\frac{t}{n}x_1}\right) E\left(e^{\frac{t}{n}x_2}\right) \dots E\left(E^{\frac{t}{n}x_n}\right)$$

$$= m_{x1}\left(\frac{t}{n}\right) m_{x2} \dots . m_{xn}\left(\frac{t}{n}\right)$$

$$= EXP\left\{\mu\frac{t}{n} + \frac{1}{2n^2}\sigma^2 r^2\right\}. EXP\left\{\mu\frac{t}{n} + \frac{1}{2n^2}\sigma^2 r^2\right\} \dots\dots EXP$$

$=\text{EXP}\left\{n\,\mu\frac{t}{n}+\frac{1}{2n^2}\sigma^2\,r^2\right\}=\text{EXP}\left\{\mu\,t+\frac{1}{2n}\sigma^2\,r^2\right\}$

That is a moment generating function of $N\left(\mu,\frac{\sigma^2}{n}\right)$, Thus

$\bar{X}\sim N\left(\mu\frac{\sigma^2}{n}\right)$.

That is

$f(\bar{X})=\sqrt{\frac{n}{2n}}\frac{1}{\sigma}e^{\frac{n(\bar{x}-\mu)^2}{2\sigma 2}}$ **;-∞<$\bar{X}$<∞**

To drive the mean and var . of $\bar{X}$:

$\text{E}(\bar{x})=\text{E}\left(\frac{1}{n}\sum_{i=1}^{n}x_1\right)=\frac{1}{n}\left(E(x_1)+E(x_2)\ldots\ldots+E(X_N)\right)$

$=\frac{1}{n}(\mu+\mu\ldots\ldots+\mu)=\frac{1}{n}n\mu=\mu$

$\therefore E(\bar{X})=\mu$

$\text{Var }(\bar{x})=\left(\frac{1}{n}\sum_{i=1}^{n}x_1\right)=\frac{1}{n^2}(\text{var}(x_1)+\text{var}(x_2)\ldots\ldots+\text{var}(x_n))$

$=\frac{1}{n}(\sigma^2+\sigma^2\ldots\ldots+\sigma^2)$

$=\frac{1}{n^2}n\sigma^2=\frac{1}{n}\sigma^2$

$\therefore\text{ var }(\bar{x})=\frac{1}{n}\sigma^2$

Example: Let $x_1,x_{2\,,\ldots\ldots and}\;x_n$ are independent and identically distributed $G(a\,,\beta)$, Find The Distribution of $\bar{X}$?

Solution:

$m_x\,(\text{t})=(1-\beta t)^{-x}$

$m_{\bar{x}}\,(\text{t})=\text{E}\left(e^{2^{\bar{x}}}\right)=\text{E}\left(e^{\frac{t}{n}(x_1+x_2+\ldots\ldots+x_n)}\right)=\text{E}\left(e^{\frac{t}{n}x_1+\frac{t}{n}x_2+\cdots+\frac{t}{n}x_n}\right)$

$= E\left(e^{\frac{t}{n}x_1}\right)E\left(e^{\frac{t}{n}x_2}\right)\ldots\ldots E\left(e^{\frac{t}{n}x_n}\right)$

$= m_{x_1}\left(\frac{t}{n}\right) m_{x_2}\left(\frac{t}{n}\right)\ldots\ldots m_{x_n}\left(\frac{t}{n}\right)$

$= \left(1-\beta\frac{t}{n}\right)^{-x}\left(1-\beta\frac{t}{n}\right)^{-x}\ldots\ldots\left(1-\beta\frac{t}{n}\right)^{-x} \Longrightarrow \left(1-\frac{\beta}{n}t\right)^{-n\alpha}$

That is a moment generating function of $G\left(n\alpha,\frac{\beta}{n}\right)$. Thus

$\bar{X} \sim G\left(n\alpha,\frac{\beta}{n}\right)$

That is

$f(\bar{x})=\frac{n^{\alpha}}{\beta^{\alpha}\,\Gamma(n\alpha)}(\bar{X})^{n\alpha-1}\, e^{\frac{n\bar{x}}{\beta}}$; $0<\bar{X}<\infty$

To drive the mean and var . of $\bar{X}$:

$E(\bar{X})=E\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)=\frac{1}{n}\left(E(x_1)+E(x_2)+\cdots+E(x_n)\right)$

$=\frac{1}{n}(\alpha\beta+\alpha\beta+\cdots+\alpha\beta)=\frac{1}{n}n\alpha\beta=\alpha\beta$

$\therefore E(\bar{X})=\alpha\beta$

$\mathrm{var}(\bar{X})=\mathrm{var}\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)=\frac{1}{n^2}\left(\mathrm{var}(x_1)+\mathrm{var}(x_2)+\cdots+\mathrm{var}(x_n)\right)=\frac{1}{n}$
$(\alpha\beta^2+\alpha\beta^2+\cdots+\alpha\beta^2)$

$=\frac{1}{n^2}n\alpha\beta^2=\frac{1}{n}\alpha\beta^2$

$\therefore (\bar{X})=\frac{1}{n}\alpha\beta^2$

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور

فراس شاكر محمود

The Distribution of S^2

Theorem:- Let $X_1 . X_2 \dots . X_n$ be observations of a random sample of size n from the normal distribution $N(\mu, \sigma^2)$ Then the sample mean .

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

and the sample variance

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

are independent and

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2}$$

Proof :- we are not prepared to prove the independence of $\overline{X}$ and S^2 at this time, so we accept it without proof here . To prove the second part . note that

$$w = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 = \sum_{i=1}^{n}\left[\frac{(X_i-\overline{X})+(\overline{X}-\mu)}{\sigma}\right]^2$$

$$= \sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2 + \frac{n(\overline{X}-\mu)^2}{\sigma^2}$$

because the cross-product term is equal to

$$2\sum_1^n \frac{(\overline{X}-\mu)(X_i-\overline{X})}{\sigma^2} = \frac{2(\overline{X}-\mu)}{\sigma^2}\sum_1^n(X_i - \overline{X}) = 0$$

But $Y_i = \frac{(\overline{X}-\mu)}{\sigma^2}, i = 1.2.3 \dots . n$

are standardized normal variables that are independent. Hence $w = \sum_1^n Y_i^2$ is $\chi^2(n)$ by corollary 5.4-3 Moreover

since $\overline{X}$ is $N(\mu, \frac{\sigma^2}{n})$ it follows that

$$Z^2 = \left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2 = \frac{n(\overline{X}-\mu)^2}{\sigma^2}$$

is $\chi^2(1)$ by Theorem 3.3-2 In this notation. Equation 5.5-1becomes

$$w = \frac{(n-1)s^2}{\sigma^2} + z^2$$

However from the face that $\overline{X}$ and S^2 are independent it follows thatZ^2 and S^2 are also independent In the mgf of W this independence permits us to write

$$E[e^{tw}] = E\left[e^{t(\frac{(n-1)s^2}{\sigma^2}+z^2)}\right] = E\left[e^{t(\frac{(n-1)s^2}{\sigma^2})}e^{tZ^2}\right] = E\left[e^{t(\frac{(n-1)s^2}{\sigma^2})}\right] E[e^{tZ^2}] .$$

Since W and z^2 have chi-square distribution we can substitute their mgfs to obtain $(1-2t)^{-n/2} = E\left[e^{t(\frac{(n-1)s^2}{\sigma^2})}\right](1-2t)^{-1/2}$

Equivalently we have $E\left[e^{t(\frac{(n-1)s^2}{\sigma^2})}\right] = (1-2t)^{-(n-1)/2}$ $t<1/2$

This of course is the mgf of a $\chi^2(n-1)$ variable accordingly($\frac{(n-1)s^2}{\sigma^2}$ has that distribution

Example:- If $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ Show that $Z=\left[\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right] \sim N(0,1)$

Solution:

Since $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$$f(\bar{X}) = \frac{1}{\frac{\sigma^2}{n}\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{(\bar{X}-\mu)^2}{\frac{\sigma^2}{n}}\right)} \quad -\infty < \bar{X} < \infty$$

$$f(\bar{X}) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{(\bar{X}-\mu)^2}{\frac{\sigma}{\sqrt{n}}}\right)} \quad -\infty < \bar{X} < \infty$$

$$Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

$$M_z(t) = E(e^{tz})$$

$$= E\left(e^{t\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)}\right)$$

$$E\left(e^{t\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)}\right) = \int_{-\infty}^{\infty} e^{t\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)} . \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \, e^{\frac{-1}{2}\left(\frac{(\bar{X}-\mu)^2}{\frac{\sigma^2}{n}}\right)} d\bar{X}$$

$$\text{let}\left[y = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right] \rightarrow \frac{\sigma y}{\sqrt{n}} = \bar{X} - \mu$$

$$\bar{X} = \frac{\sigma y}{\sqrt{n}} + \mu \rightarrow d\bar{X} = \frac{\sigma}{\sqrt{n}} dy$$

$$E(e^{tz}) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)} . \, e^{\frac{-1}{2}\left(\frac{(\bar{X}-\mu)^2}{\frac{\sigma^2}{n}}\right)} d\bar{X}$$

$$E(e^{tz}) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ty} \, e^{\frac{-1}{2}y} \frac{\sigma}{\sqrt{n}} dy$$

$$E(e^{tz}) = \frac{\sigma}{\sqrt{n}} \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(\frac{y^2-2ty}{2})} dy$$

$$E(e^{tz}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(\frac{y^2-2t+t^2-t^2}{2})} dy$$

$$E(e^{tz}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{y^2-2t+t^2}{2}\right)-\frac{t^2}{2}} dy$$

$$E(e^{tz}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(y-t)^2}{2}} \, e^{\frac{t^2}{2}} dy$$

$$E(e^{tz}) = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(\frac{(y-t)^2}{2})} dy$$

Let h=y-t → dh = dy

$$E(e^{tz}) = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2}h^2}\, dh$$

$$E(e^{tz}) = e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{\frac{-1}{2}h^2}\, dh \;\rightarrow\; \frac{1}{\sqrt{2\pi}}\, e^{\frac{-1}{2}h^2} = 1 \sim N(0,1)$$

$$E(e^{tz}) = e^{\frac{t^2}{2}} \sim N(0,1)$$

$$Z = \left[\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right] \sim N(0,1)$$

Student t-distribution:-

Theorem :-Let $T = \frac{Z}{\sqrt{\frac{U}{r}}}$

where Z is a random variable that is $N(0,1)$, U is a random variable that is $X^2(r)$ and Z and U are independent . Then T has a t distribution with pdf $f(t) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi r}\,\Gamma\left(\frac{r}{2}\right)} \frac{1}{\left(1+\frac{t^2}{r}\right)^{\frac{r+1}{2}}} \quad -\infty < t < \infty$

proof :- The joint pdf of Z and U is

$$g(z,u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{\frac{r}{2}}} u^{\frac{r}{(2-1)}} e^{-\frac{u}{2}}$$

the cdf $F(t)=P(T\leq t)$ of Tis given by

$$F(t) = P\left(\frac{Z}{\sqrt{\frac{U}{r}}} \leq t\right)$$

$$= P\left(Z \leq \sqrt{\frac{U}{rt}}\right)$$

$$= \int_0^{\infty} \int_{-\infty}^{\sqrt{\left(\frac{u}{r}\right)}t} g(z,u)\, dz\, du.$$ That is $F(t) = \frac{1}{\sqrt{\pi}\,\Gamma\left(\frac{r}{2}\right)} \int_0^{\infty} \left[\int_{-\infty}^{\sqrt{\left(\frac{u}{r}\right)}t} \frac{e^{-\frac{z^2}{2}}}{\frac{2(r+1)}{2}} dz\right] u^{\frac{r}{2}} e^{-\frac{u}{2}}\, du$

the pdf of T is the derivative of the cdf , so, applying the fundamental theorem of calculus to the inner integral

we find that $f(t) = F\'(t) = \frac{1}{\sqrt{\pi}\,\Gamma\left(\frac{r}{2}\right)} \int_0^{\infty} \frac{e^{-\left(\frac{u}{2}\right)\left(\frac{t^2}{r}\right)}}{2^{(r+1)/2}} \sqrt{\frac{u}{r}}\, u^{\frac{r}{2}-1} e^{-u/2} du$

$$= \frac{1}{\sqrt{\pi}\,\Gamma\left(\frac{r}{2}\right)} \int_0^{\infty} \frac{u^{\frac{r+1}{2}-1}}{2^{(r+1)/2}} e^{-\left(\frac{u}{2}\right)\left(\frac{1+t^2}{r}\right)} du$$

In the integral, make the change of variables $y=(1+t^2/r)u$, so that $\frac{du}{dy} = \frac{1}{1+t^2/r}$

Thus, $f(t) = \frac{\Gamma[(r+1)]/2}{\sqrt{\pi r}\Gamma(\frac{r}{2})}\left[\frac{1}{(1+t^2/r)^{(r+1)/2}}\right]\int_0^\infty \frac{y^{(r+1)/(2-1)}}{\Gamma[\frac{r+1}{2}]2^{(r+1)/2}} e^{-y/2}dy$

The integral in this last expression for f(t) is equal to 1 because the integrand is like the pdf of a chi-square distribution with r+1 degrees of freedom . Hence , the pdf is

$$f(t) = \frac{\Gamma[^{(r+1)}/_2]}{\sqrt{\pi r}\Gamma\left(\frac{r}{2}\right)}\frac{1}{\left(1+\frac{t^2}{r}\right)^{\frac{r+1}{2}}} \quad -\infty < t < \infty$$

Example: if T ~t(10) then what is the probability that T is at least 2.228?

Solution:

$P(T \geq 2\cdot 228) = 1 - P(T < 2\cdot 228)$

$= 1 - 0\cdot 975$ (from t- table)

$= 0\cdot 025$

The F-distribution

Next Consider two independent chi- square random variables U and V having and in degrees of freedoms respectively. The joint p d f h(u ,v)

of u and v is then

$$h(u,v)=\begin{cases} \frac{1}{\Gamma(^{r_1}/_2)\Gamma(^{r_2}/_2)\, 2^{^{r_1+r_2}/_2}}\, u^{^{r_1}/_2-1}\, v^{^{r_2}/_2-1}\, e^{^{-(u+v)}/_2} \\ 0 & 0 < u, v < \infty \end{cases}$$

we define the new random variable w=$\frac{^v/_{r_1}}{^v/_{r_2}}$ and we propose finding the p d f $g_1(w)$of w, z=v then

$$w=\frac{^u/_{r_1}}{^v/_{r_2}}$$

define a one to one transformation that maps the set S={(u ,v):<u<∞, 0<z<∞} onto the T= {(w, z):, 0 < w < ∞, 0 < z < ∞} since u=$(\frac{r_1}{r_2})$zw, v=z the absolute value of the Jacobean of to the transformation is $|J| = \left(^{r_1}/_{r_2}\right)z$ the joint 2 p d f g(w ,z)of the random variables w and z=v is them

$$G(w,z)=\frac{1}{\Gamma(^{r_1}/_2)\Gamma(^{r_2}/_2)2^{^{r_1+r_2}/_2}}\left(\frac{r_1 zw}{r^2}\right)^{\frac{r_1-2}{2}} z^{\frac{(r_2-2)}{2}} exp[\frac{-z}{2}\left(\frac{r_1 w}{r_2}+1\right)]\frac{r_1 z}{r_2}$$

provided that (w, z) ∈ T and zero elsewhere. The marginal p d f $g_1(w)$) of w is then

$$g_1(w) = \int_{-\infty}^{\infty} g(w,z)dz$$

$$\int_0^\infty \frac{(^{r_1}/_{r_2})^{^{r_1}/_2} w^{^{r_1}/_2-1}}{\Gamma\left(^{r_1}/_2\right)\Gamma(^{r_2}/_2)2^{^{r_1+r_2}/_2}} z^{^{(r_1+r_2)}/_2-1} \exp[\frac{-z}{2}(\frac{r_1 w}{r_2}+1)]dz$$

If we change the variable of integration by writing $Y=\frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)$. It can be seen that

$$g_1(w) = \int_0^\infty \frac{r_1/r_2^{\,r_1/2}\, w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{2y}{r_1 w/r_2+1}\right)^{r_1+r_2/2-1} * e^{-y} * \frac{2}{r_1 w/r_2 + 1} dy$$

$$= \begin{cases} \frac{\Gamma[(r_1+r_2/2](r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)} \cdot \frac{w^{n/2-1}}{(1+r_1 w/r_2)(r_1+r_2)/2} & 0 < w < \infty \\ 0 & 0.w \end{cases}$$

Accordingly, if U and V are independent chi Square variable with $r_1$and r_2 degrees of freedom, respectively, then w= $(U/r_1)/(V/r_2)$ has the p d f $g_1(w)$the distribution of this nandam variable is usually called an F-distribution and we often call ration which we have denoted by w ,f. That is,F=$\frac{U/r_1}{V/r_2}$.

Example

Let F have an F-distribution with $r_1$and r_2 degrees of freedom, we can write F=$(r_1/r_2)(U/V)$ where U and V are independent X^2 random Variable with $r_1$and $r_2$degrees of freedom respectively

Hence for the kth moment of F, by independence we have E$(F^K) = \frac{r_2{}^k}{r_1}\ E(U^k)E(V^{-k})$. Provided of course that both expectations on the night side exist K> $(r_1/2)$ is always true, the first expectation always exists. The second expectation, however, exists if r_2> 2k. i.e. the denominator degrees of freedom must exceed twice k Assuming this is true , it follows that the mean of f-is given by

E(F)=$\frac{r_2}{r_1} r_1 \frac{2^{-1}\Gamma(\frac{r_2}{2}-1)}{\Gamma(\frac{r_2}{2})} = \frac{r_2}{r_2-2}$

Theorem 1:

Let U and V be two independent random variables having chi-squared distributions with v_1 and v_2 degrees of freedom, respectively. Then the distribution of the random variable $F = \frac{U/v_1}{V/v_2}$ is given by the density function

$$h(f) = \begin{cases} \frac{\Gamma[(v_1+v_2)/2](v_1/v_2)^{v_1/2}}{\Gamma(v_1/2)\Gamma(v_2/2)} \frac{f^{(v_1/2)-1}}{(1+v_1 f/v_2)^{(v_1+v_2)/2}}, & f > 0, \\ 0, & f \le 0. \end{cases}$$

This is known as the ***F*-distribution** with v_1 and v_2 degrees of freedom (d.f.).

the density function will not be used and is given only for completeness. The curve of the *F*-distribution depends not only on the two parameters v_1 and v_2 but also on the order in which we state them. Once these two values are given, we can identify the curve. Typical *F*-distributions are shown in Figure 1.

Let f_α be the f-value above which we find an area equal to α. This is illustrated by the shaded region in Figure 2. Hence, the f-value with 6 and 10 degrees of freedom, leaving an area of 0.05 to the right, is $f_{0.05} = 3.22$. By means of the following theorem.

Theorem2:

Writing $f_\alpha(v_1, v_2)$ for f_α with v_1 and v_2 degrees of freedom, we obtain

$$f_{1-\alpha}(v_1, v_2) = \frac{1}{f_\alpha(v_2, v_1)}.$$

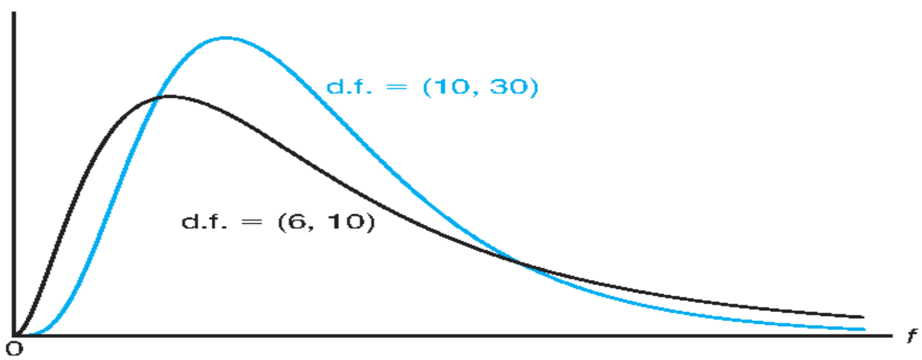


Figure1: Typical F-distributions.

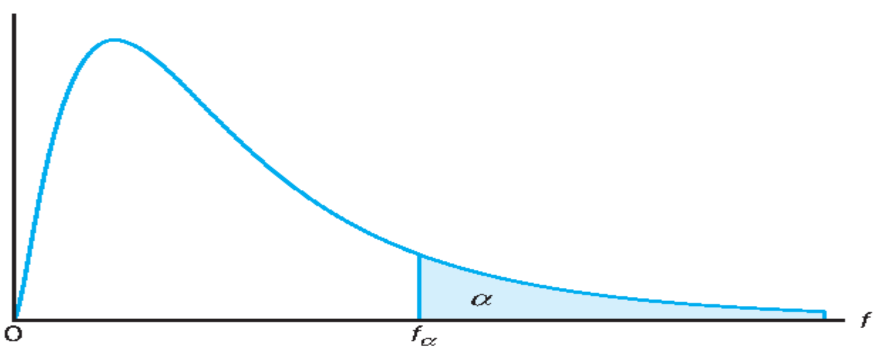


Figure2: Illustration of the f_α for the F-distribution.

Thus, the *f*-value with 6 and 10 degrees of freedom, leaving an area of 0.95 to the right, is

$$f_{0.95}(6,10) = \frac{1}{f_{0.05}(10,6)} = \frac{1}{4.06} = 0.246.$$

The F-Distribution with Two Sample Variances

Suppose that random samples of size n_1 and n_2 are selected from two normal populations with variances σ_1^2 and σ_2^2, respectively. From Theorem 8.4, we know that

$$\chi_1^2 = \frac{(n_1-1)S_1^2}{\sigma_1^2} \text{ and } \chi_2^2 = \frac{(n_2-1)S_2^2}{\sigma_2^2}$$

are random variables having chi-squared distributions with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom. Furthermore, since the samples are selected at random, we are dealing with independent random variables. Then, using Theorem 1 with $\chi_1^2 = U$ and $\chi_2^2 = V$, we obtain the following result.

Theorem 3 :

If S_1^2 and S_2^2 are the variances of independent random samples of size n_1 and n_2 taken from normal populations with variances σ_1^2 and σ_2^2, respectively, then

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$$

has an *F*-distribution with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom.

Central limit theorem

Def. If $\bar{X}$ is the mean of random sample ; $x_1, x_2, \dots, x_n$ of size n from a distribution . with finite mean and finite variance then the distribution of r.v. $w = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim N(0,1)$

in the limit as $\lim_{n\to\infty}\left[\frac{\bar{X}-\mu}{s/\sqrt{n}}\right] \sim N(0,1)$

$\lim_{n\to\infty} p\{|y_n - c| < \epsilon\} = 1$ or $\lim_{n\to\infty} p\{|y_n - c| \geq \epsilon\} = 0$

Chebyshev's inequality

Where $\lim_{n\to\infty} p\{|y_n - c| < \epsilon\} = 1 - \frac{1}{k^2}$ lower

$\lim_{n\to\infty} p\{|y_n - c| \geq \epsilon\} = \frac{1}{k^2}$ uper

Example: let $\bar{X}_n$ denoted the mean of a r.s. of size n from distribution having the mean μ and the variance σ^2 show that $\bar{X}_n \xrightarrow{c.s} \mu$.

Solution:

$p[|\bar{X}_n - \mu| < \epsilon] \geq 1 - \frac{1}{k^2}$

$\lim_{n\to\infty} p[|\bar{X}_n - \mu| < \epsilon] = 1 \qquad \forall\epsilon > 0$

Since a distribution is $\bar{X}_n$

$\therefore mean(\bar{X}_n) = \mu\,, var(\bar{X}_n\,) = \frac{\sigma^2}{n} \Rightarrow S.D. = \sqrt{var(\bar{X}_n)} = \frac{\sigma}{\sqrt{n}}$

Let $\epsilon = k\frac{\sigma}{\sqrt{n}} \Rightarrow k = \frac{\epsilon\sqrt{n}}{\sigma}$

$\lim_{n\to\infty} p\left[|\bar{X}_n - \mu| < k\frac{\sigma}{\sqrt{n}}\right] \geq \lim_{n\to\infty}\left[1 - \frac{1}{(\frac{\epsilon\sqrt{n}}{\sigma})^2}\right] = 1$

$\therefore \lim_{n\to\infty} p[|\bar{X}_n - \mu| < \epsilon] = 1$

$\therefore \bar{X}_n \xrightarrow{c.s} \mu$

Example: Show that $\frac{s^2}{n-1} \xrightarrow{c\,.\,s} \sigma^2$

Solution:

$p[|y_n - c| < \epsilon] \geq 1 - \frac{1}{k^2}$

$\lim_{n\to\infty} p[|y_n - c| < \epsilon] = 1 \;\; ; \; \forall \epsilon > 0$

Since a distribution is $\frac{s^2}{n-1} \sim {\chi^2}_{(n-1)}$

$\therefore mean = (n-1) \;\; , var = 2(n-1)$

$p\left[\left|\frac{s^2}{n-1} - \sigma^2\right| < \epsilon\right] \geq 1 - \frac{1}{k^2}$ $\qquad\qquad * \{\frac{n-1}{\sigma^2}\}$

$p\left[\left|\frac{s^2}{\sigma^2} - (n-1)\right| < \frac{\epsilon(n-1)}{\sigma^2}\right] \geq 1 - \frac{1}{k^2}$

$let \; \frac{\epsilon(n-1)}{\sigma^2} = k\sqrt{2(n-1)}$

$\Rightarrow k = \frac{\epsilon(n-1)}{\sigma^2\sqrt{2(n-1)}} \Rightarrow k^2 = \frac{\epsilon^2(n-1)^2}{\sigma^4 2(n-1)}$

$\lim_{n\to\infty} p\left[\left|\frac{s^2}{\sigma^2} - (n-1)\right| < k\sqrt{2(n-1)}\right] \geq \lim_{n\to\infty}\left[1 - \frac{1}{\frac{\epsilon^2(n-1)^2}{\sigma^4 2(n-1)}}\right]$

$\lim_{n\to\infty} p\left[\left|\frac{s^2}{\sigma^2} - (n-1)\right| < k\sqrt{2(n-1)}\right] \geq \lim_{n\to\infty}\left[1 - \frac{2\sigma^4}{\epsilon^2(n-1)}\right] = 1$

$\therefore \lim_{n\to\infty} p\left[\left|\frac{s^2}{n-1} - \sigma^2\right| < \epsilon\right] = 1$

$\Rightarrow \frac{s^2}{n-1} \xrightarrow{c\,.\,s} \sigma^2$

Example: If $x_n \xrightarrow{c\,.\,s} c$ show that $\sqrt{x_n} \xrightarrow{c\,.\,s} \sqrt{c}$

Solution:

$\lim_{n\to\infty}[|x_n - c| < \epsilon] = \lim_{n\to\infty}\left[\left|(\sqrt{x_n} - \sqrt{c})(\sqrt{(x_n} + \sqrt{c})\right| < \epsilon\right]$

$= \lim_{n\to\infty}\left[\left|\sqrt{x_n} - \sqrt{c}\right|\left|\sqrt{x_n} + \sqrt{c}\right| < \epsilon\right] \qquad\qquad \div \left|\sqrt{x_n} + \sqrt{c}\right|$

$$= \lim_{n\to\infty}\left[\left|\sqrt{x_n}-\sqrt{c}\right| < \frac{\epsilon}{\left|\sqrt{x_n}+\sqrt{c}\right|}\right]$$

$$let\ \frac{\epsilon}{\left|\sqrt{x_n}+\sqrt{c}\right|} = \epsilon'$$

$$= \lim_{n\to\infty}\left[\left|\sqrt{x_n}-\sqrt{c}\right| < \epsilon'\right]$$

$$\lim_{n\to\infty}[|x_n - c| < \epsilon] = \lim_{n\to\infty}\left[\left|\sqrt{x_n}-\sqrt{c}\right| < \epsilon'\right]$$

$$Since\ x_n \xrightarrow{c\,.\,s} c \Rightarrow \lim_{n\to\infty}[|x_n - c| < \epsilon] = 1$$

$$\therefore \lim_{n\to\infty}\left[\left|\sqrt{x_n}-\sqrt{c}\right| < \epsilon'\right] = 1$$

$$\therefore \sqrt{x_n} \xrightarrow{c\,.\,s} \sqrt{c}\ .$$

Example: Let w_n denote a random variable with mean μ and variance $\frac{b}{n^p}$, where $p > 0$, μ and b are constants (not functions of n). Prove that w_n converges to μ . or ($w_n \xrightarrow{c\,.s} \mu$)

Solution:

$$p[|w_n - \mu| < \epsilon = k\sigma] \geq 1 - \frac{1}{k^2}$$

$$\text{since mean} = \mu \text{ and variance} = \frac{b}{n^p}$$

$$let\ \epsilon = k\sqrt{\frac{b}{n^p}}$$

$$k = \frac{\epsilon}{\sqrt{b}/\sqrt{n^p}}$$

$$k = \frac{\epsilon\sqrt{n^p}}{\sqrt{b}}$$

$$\lim_{n\to\infty} p\left[|w_n - \mu| < k\sqrt{\frac{b}{n^p}}\right] \geq \lim_{n\to\infty}\left[1 - \frac{1}{\left(\frac{\epsilon\sqrt{n^p}}{\sqrt{b}}\right)^2}\right]$$

$$\lim_{n\to\infty} p\left[|w_n-\mu|<k\sqrt{\frac{b}{n^p}}\right] \geq \lim_{n\to\infty}\left[1-\frac{b}{\epsilon^2 n^p}\right]=1$$

$$\lim_{n\to\infty} p\left[|w_n-\mu|<k\sqrt{\frac{b}{n^p}}\right]=1$$

$$\lim_{n\to\infty} p[|w_n-\mu|<\epsilon]=1$$

$$w_n \xrightarrow{c.s} \mu$$

Example: Let the random variable Y_n have a distribution that is $b(n,p)$

i. Prove that ${Y_n}/{n} \xrightarrow{c.s} p$.

ii. Prove that $1-{Y_n}/{n} \xrightarrow{c.s} 1-p$.

Solution:

i. $p\left[\left|{Y_n}/{n}-p\right|<\epsilon\right] \geq 1-\frac{1}{k^2}$

Since Y_n have a distribution that is $b(n,p) \Rightarrow$

Mean $(Y_n)=np$ and

Var $(Y_n)=npq$

$$p\left[\left|{Y_n}/{n}-p\right|<\epsilon\right] \geq 1-\frac{1}{k^2} \qquad *n$$

$$p[|Y_n-np|<n\epsilon] \geq 1-\frac{1}{k^2}$$

Let $n\epsilon = k\sqrt{npq}$

$$k=\frac{n\epsilon}{\sqrt{npq}}$$

$$\lim_{n\to\infty} p[|Y_n-np|<k\sqrt{npq}] \geq \lim_{n\to\infty}\left[1-\frac{1}{(\frac{n\epsilon}{\sqrt{npq}})^2}\right]$$

$$\lim_{n\to\infty} p[|Y_n-np|<k\sqrt{npq}] \geq \lim_{n\to\infty}\left[1-\frac{1}{\frac{n^2\epsilon^2}{npq}}\right]$$

$$\lim_{n\to\infty} p\left[|Y_n - np| < k\sqrt{npq}\right] \geq \lim_{n\to\infty}\left[1 - \frac{pq}{n\epsilon^2}\right] = 1$$

$$\lim_{n\to\infty} p\left[|Y_n - np| < k\sqrt{npq}\right] = 1$$

$$\lim_{n\to\infty} p\left[\left|{}^{Y_n}/_{n} - p\right| < \epsilon\right] = 1$$

$${}^{Y_n}/_{n} \xrightarrow{c.s} p$$

ii. $$p\left[\left|1 - {}^{Y_n}/_{n} - (1-p)\right| < \epsilon\right] \geq 1 - \frac{1}{k^2}$$

$$p\left[\left|1 - {}^{Y_n}/_{n} - 1 + p\right| < \epsilon\right] \geq 1 - \frac{1}{k^2}$$

$$p\left[\left|- {}^{Y_n}/_{n} + p\right| < \epsilon\right] \geq 1 - \frac{1}{k^2}$$

$$p\left[|-1| . \left|{}^{Y_n}/_{n} - p\right| < \epsilon\right] \geq 1 - \frac{1}{k^2}$$

$$p\left[\left|{}^{Y_n}/_{n} - p\right| < \epsilon\right] \geq 1 - \frac{1}{k^2} \qquad * n$$

$$p[|Y_n - np| < n\epsilon] \geq 1 - \frac{1}{k^2}$$

$$p[|Y_n - np| < n\epsilon] \geq 1 - \frac{1}{k^2}$$

Let $n\epsilon = k\sqrt{npq}$

$$k = \frac{n\epsilon}{\sqrt{npq}}$$

$$p\left[|Y_n - np| < k\sqrt{npq}\right] \geq 1 - \frac{1}{k^2}$$

$$\lim_{n\to\infty} p\left[|Y_n - np| < k\sqrt{npq}\right] \geq \lim_{n\to\infty}\left[1 - \frac{1}{\left(\frac{n\epsilon}{\sqrt{npq}}\right)^2}\right]$$

$$\lim_{n\to\infty} p\left[|Y_n - np| < k\sqrt{npq}\right] \geq \lim_{n\to\infty}\left[1 - \frac{1}{\frac{n^2\epsilon^2}{npq}}\right]$$

$$\lim_{n\to\infty} p\left[|Y_n - np| < k\sqrt{npq}\right] \geq \lim_{n\to\infty}\left[1 - \frac{pq}{n\epsilon^2}\right] = 1$$

$$\lim_{n\to\infty} p\left[|Y_n - np| < k\sqrt{npq}\right] = 1$$

$$\lim_{n\to\infty} p\left[\left|1 - {}^{Y_n}\!/_{n} - (1-p)\right| < \epsilon\right] = 1$$

$$1 - {}^{Y_n}\!/_{n} \xrightarrow{c.s} 1 - p\,.$$

Example: Let $x_1, x_2, \dots, x_{25}$ be a r.v. of size 25 $\sim N(75{,}100)$ compute $p(71 < \bar{X} < 79)$

Solution:

$$\bar{X} \sim N\left(\mu, \frac{s^2}{n}\right) \Rightarrow \bar{X} \sim N(75, \frac{100}{25})$$

$$p(71 < \bar{X} < 79) \Rightarrow p\left(\frac{71-\mu}{{}^{s}\!/_{\sqrt{n}}} < \frac{\bar{X}-\mu}{{}^{s}\!/_{\sqrt{n}}} < \frac{79-\mu}{{}^{s}\!/_{\sqrt{n}}}\right)$$

$$= p\left(\frac{71-75}{2} < Z < \frac{79-75}{2}\right) = p(-2 < z < 2)$$

$$= F(2) - F(-2) = F(2) - [1 - F(2)]$$

$$= 2F(2) - 1 = 0.954 \qquad \textit{because } F(2) = 0.977$$

$$p(71 < \bar{X} < 79) = 0.954$$

Example: If $\bar{X}$ is the mean of random sample of size n from a normal distribution with mean μ and variance 100 find n sample size where $p(\mu - 5 < \bar{X} < \mu + 5) = 0.954$

Solution:

$$p\left(\frac{(\mu-5)-\mu}{{}^{\sigma}\!/_{\sqrt{n}}} < \frac{(\bar{X}-\mu)}{{}^{\sigma}\!/_{\sqrt{n}}} < \frac{(\mu+5)-\mu}{{}^{\sigma}\!/_{\sqrt{n}}}\right) = 0.954$$

$$p\left(\frac{(\mu-5-\mu)\sqrt{n}}{10} < z < \frac{(\mu+5-\mu)\sqrt{n}}{10}\right) = 0.954$$

$$p\left(\frac{-5\sqrt{n}}{10} < z < \frac{5\sqrt{n}}{10}\right) = 0.954$$

$$p\left(\frac{-\sqrt{n}}{2} < z < \frac{\sqrt{n}}{2}\right) = 0.954$$

$$p\left(z < \frac{\sqrt{n}}{2}\right) - \mathrm{p}\left(\mathrm{z} > \frac{-\sqrt{n}}{2}\right) = 0.954$$

$$p\left(z < \frac{\sqrt{n}}{2}\right) - \left(1 - p\left(z < \frac{\sqrt{n}}{2}\right)\right) = 0.954$$

$$2p\left(z < \frac{\sqrt{n}}{2}\right) - 1 = 0.954$$

$$2p\left(z < \frac{\sqrt{n}}{2}\right) = 0.954 + 1$$

$$2p\left(z < \frac{\sqrt{n}}{2}\right) = 1.954 \qquad \div 2$$

$$p\left(z < \frac{\sqrt{n}}{2}\right) = 0.977$$

$$F(2) = 0.977$$

$$\therefore \frac{\sqrt{n}}{2} = 2 \Rightarrow \sqrt{n} = 4$$

$$\Rightarrow n = 16$$

Example: Let $x_{1,}x_2, \dots, x_{25}$ and $y_{1,}y_{2,\dots,}y_{25}$ be two random samples from two independent normal distribution N(0,16) , N(1,9) respectively let $\bar{x}$ *and* $\bar{y}$ denote the corresponding sample means compute $p(\bar{x} > \bar{y})$

Solution:

$$\text{since } \mathrm{x_i} \sim \mathrm{N}(0{,}16) \Rightarrow \bar{\mathrm{x}} \sim \mathrm{N}\left(\mu, \frac{\mathrm{s}^2}{\mathrm{n}}\right) = (0, \frac{16}{25})$$

$$\text{since } \mathrm{y} \sim \mathrm{N}(1{,}9) \Rightarrow \bar{\mathrm{y}} \sim \mathrm{N}\left(\mu, \frac{\mathrm{s}^2}{\mathrm{n}}\right) = (1, \frac{9}{25})$$

$$p(\bar{x} > \bar{y}) = p(\bar{x} - \bar{y} > 0)$$

$$\bar{x} - \bar{y} \sim \left(0 - 1, \frac{9+16}{25}\right) \Rightarrow \bar{x} - \bar{y} \sim (-1{,}1)$$

$p(\bar{x}-\bar{y}>0)=p[\frac{(\bar{x}-\bar{y})-(-1)}{1}>\frac{0-(-1)}{1}]$

$=p(z>1)$

$=1-p(z\leq 1)$

$p(z\leq 1)=0.8438$

$=1-0.8438=0.1562$

Example: Compute an approximate prove that :- the mean at r.s of size 15 from a distribution having $f(x)=3x^2\ ; 0<x<1$ is between $\frac{3}{5}$ and $\frac{4}{5}$?

Solution:

$$f(x)=\begin{cases}3x^2\ , & 0<x<1\\ 0 & , o.w\end{cases}$$

$$E(x)=\int_0^1 xf(x)dx=\int_0^1 3x^3dx$$

$$=\left[\frac{3}{4}x^4\right]_0^1=\frac{3}{4}$$

$$E(x^2)=\int_0^1 x^2f(x)dx=\int_0^1 3x^4dx$$

$$=\left[\frac{3}{5}x^5\right]_0^1=\frac{3}{5}$$

$$var(x)=E(x^2)-[E(x)]^2$$

$$var(x)=\frac{3}{5}-\left(\frac{3}{4}\right)^2=\frac{3}{5}-\frac{9}{16}=\frac{3}{80}$$

$$\bar{x}\sim\left(\mu,\frac{\sigma^2}{n}\right)\Rightarrow\bar{x}\sim\left(\frac{3}{4},\frac{3/80}{15}\right)\sim\left(\frac{3}{4},\frac{1}{400}\right)$$

$$\therefore\ \mu=\frac{3}{4}\ \ and\ var(\bar{x})=\frac{1}{400}\Rightarrow\sqrt{var(\bar{x})}=\frac{1}{20}$$

$$p\left(\frac{3}{5}<\bar{x}<\frac{4}{5}\right)=p\left(\frac{\frac{3}{5}-\mu}{1/20}<z<\frac{\frac{4}{5}-\mu}{1/20}\right)$$

$$p\left(\frac{\frac{3}{5}-\frac{3}{4}}{1/20}<z<\frac{\frac{4}{5}-\frac{3}{4}}{1/20}\right)=p\left(\frac{\frac{12-15}{20}}{1/20}<z<\frac{\frac{16-15}{20}}{1/20}\right)$$

$= p(-3 < z < 1)$

$= p(z < 1) - p(z > -3)$

$= p(z < 1) - [1 - p(z \leq 3)]$

$= p(z < 1) + p(z \leq 3) - 1$

$= F(1) + F(3) - 1 = 0.8531 + 0.9989 - 1 = 0.852$

الجداول الاحصائية لتوزيع الطبيعي المعياري لاستخراج قيمة الدالة التوزيعية للعددين الواحد والثلاثة عند مستوى افقي *0.05*

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور فراس شاكر محمود

Limiting Moment-Generating Functions

To find the limiting distribution function of a random variable V by use of the definition of limiting distribution function obviously requires that we know F,(y) for each positive integer n. This is precisely the problem we should like to avoid. If it exists, the moment-generating function that corresponds to the distribution function F,(y) often provides a convenient method of determining the limiting distribution function. To emphasize that the distribution of a random variable Y, depends upon the positive integer n, in this lecture we shall write the moment-generating function of Y, in the form M(t; n). The following theorem, which is essentially Curtiss' modification of a theorem of Lévy and Cramér, explains how the moment-generating function may be used in problems of limiting distributions. A proof of the theorem requires a knowledge of that same facet of analysis that permitted us to assert that a moment-generating function, when it exists, uniquely determines a distribution. Accordingly, no proof of the theorem will be given.

Theorem 1. Let the random variable Y_n , have the distribution function $F_n(y)$ and the moment-generating function M(t ; n) that exists for $-h < t < h$ for all n. If there exists a distribution function F(y), with corresponding moment-generating function M(t), defined for $|t| \leq h_1 < h$, such that $\lim_{n\to\infty} M(t; n) = M(t)$, then Y_n , has a limiting distribution with distribution function F(y). Several illustrations of the use of **Theorem 1.** In some of these examples it is convenient to use a certain limit that is established in some courses in advanced calculus. We refer to a limit of the form

$$\lim_{n\to\infty}\left(1+\frac{b}{n}+\frac{\varphi(n)}{n}\right)^{cn}$$

where b and c do not depend upon n and where $\lim_{n\to\infty}(\varphi(n)) = 0$. then

$\lim_{n\to\infty}\left(1+\frac{b}{n}+\frac{\varphi(n)}{n}\right)^{cn} = \lim_{n\to\infty}\left(1+\frac{b}{n}\right)^{cn} = e^{bc}$.

Example:

$\lim_{n\to\infty}\left(1-\frac{t^2}{n}+\frac{t^3}{n^{3/2}}\right)^{-n/2} = \lim_{n\to\infty}\left(1-\frac{t^2}{n}+\frac{t^3/\sqrt{n}}{n^{3/2}}\right)^{-n/2}$

Here b= - t^2 ، c= $\frac{-1}{2}$ ، and $\varphi(n) = t^3/\sqrt{n}$

Accordingly for every fixed value of t, the limit is $e^{t^2/2}$

Theorem 2. let $Y_n \sim b(n،p)$ show that the limit of Y_n as $n \to \infty$.

Proof:

Since $Y_n \sim b(n،p)$

So $M_{Y_n}(t،n) = (pe^t + q)^n$ q=1-p

$$\mu = np \to p = \frac{\mu}{n}$$

$M_{Y_n}(t،n) = (pe^t + 1 - p)^n$ سحب p عامل مشترك

$M_{Y_n}(t،n) = (p(e^t - 1) + 1)^n$ نعوض قيمة p

$$M_{Y_n}(t،n) = \left(\frac{\mu}{n}(e^t - 1) + 1\right)^n$$

$\left(1+\frac{x}{n}\right)^n = e^x$ where $x = \mu(e^t - 1)$

$$\therefore M_{Y_n}(t,n) = \left(\frac{\mu(e^t-1)}{n} + 1\right)^n = e^{\mu(e^t-1)}$$

$Y_n = e^{\mu(e^t-1)} \qquad \sim poisson(\mu)$

$\therefore \lim_{n\to\infty} Y_n \sim poisson(\mu)$

Example: let $Z_n \sim Poisson(n)$ find the limiting distribution of $Y_n = \frac{Z_n - n}{\sqrt{n}}$?

Solution :

$$M_{Y_n}(t,n) = E(e^{Y_n}) = E\left(e^{t\frac{Z_n-n}{\sqrt{n}}}\right)$$

$$= e^{t\frac{-n}{\sqrt{n}}} \quad E\left(e^{t\frac{Z_n}{\sqrt{n}}}\right)$$

Since $Z_n \sim Poisson(n) \ \rightarrow \ M_{Z_n} = e^{n(e^t-1)}$

$$= e^{t\frac{-n}{\sqrt{n}}} \quad e^{n(e^t-1)} = e^{-t\sqrt{n}} \ e^{ne^{\frac{t}{\sqrt{n}}}-n}$$

$$M_{Y_n}(t) = e^{-\sqrt{n}\ t - n + ne^{\frac{t}{\sqrt{n}}}}$$

$$e^{\frac{t}{\sqrt{n}}} = 1 + \frac{t}{\sqrt{n}} + \frac{1}{2!}\left(\frac{t}{\sqrt{n}}\right)^2 + \cdots$$

$$M_{Y_n}(t) = e^{-\sqrt{n}\ t - n + n\left(1+\frac{t}{\sqrt{n}}+\frac{1}{2!}\left(\frac{t}{\sqrt{n}}\right)^2+\cdots\right)}$$

$$M_{Y_n}(t) = e^{-\sqrt{n}\ t - n + n + \sqrt{n}\ t + \frac{t^2}{2} + \cdots}$$

$M_{Y_n}(t) = e^{\frac{t^2}{2}+\frac{\emptyset(n)}{n}}$ $\emptyset(n) \to 0$

$M_{Y_n}(t) = e^{\frac{t^2}{2}} \sim N\left(0،1\right)$

Example: Let Y_n denote the n^{th} order statistic of r.s from a distribution of the continuous type that has distribution function $F(x)$ and p.d.f. $f(x)$ find the limiting distribution of $Z_n = n[1 - F(Y_n)]$?

Solution : Note that (Y_1 order smaller $< Y_2 < \cdots < Y_n$ order larger)

Since Y_n $\therefore$ order largest

$g(Y_n) = n[F(Y_n)]^{n-1} f(Y_n)$

Since $Z_n = n[1 - F(Y_n)]$

$Z_n = n - nF(Y_n)$

$nF(Y_n) = n - Z_n \to F(Y_n) = 1 - \frac{Z_n}{n}$

هذه الدالة التوزيعية

يجب ان نجد الدالة الاحتمالية اي اشتقاق الدالة التوزيعية بالنسبة ل Z_n

$F(Y_n)\frac{dY_n}{dZ_n} = -\frac{1}{n}$

$\frac{dY_n}{dZ_n} = \frac{-1}{nf(Y_n)} \to \left|\frac{dY_n}{dZ_n}\right| = \left|\frac{-1}{nf(Y_n)}\right| = \frac{1}{nf(Y_n)}$

لان

$$h(Z_n) = g(Y_n)|\,J\,|$$

$$h(Z_n) = n\left[1-\frac{Z_n}{n}\right]^{n-1} f(Y_n)\frac{1}{nf(Y_n)} \rightarrow h(Z_n) = \left[1-\frac{Z_n}{n}\right]^{n-1}$$

$$\lim_{n\to\infty}\left[1-\frac{Z_n}{n}\right]^{n-1} \rightarrow \lim_{n\to\infty}\left[1-\frac{Z_n}{n}\right]^{n} \lim_{n\to\infty}\left[1-\frac{Z_n}{n}\right]^{-1}$$

$$(1-0)\lim_{n\to\infty}\left[1-\frac{Z_n}{n}\right]^{n} = e^{-Z_n} \sim \text{Gamma}\left(1،1\right) \text{ or Exp }(1)$$

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

Lecture Note On Mathematical Statistics 1

B.Sc. in Mathematics

Fourth Stage

Assist. Prof. Dr. Feras Shaker Mahmood

Order Statistics

المحاضرة التاسعة

الكورس الاول

ORDER STATISTICS

In practice, the random variables of interest may depend on the relative magnitudes of the observed variable. For example, we may be interested in the maximum mileage per gallon of a particular class of cars. In this section, we study the behavior of ordering a random sample from a continuous distribution.

Definition *Let $X_1, \ldots, X_n$ be a random sample from a continuous distribution with pdf $f(x)$. Let $Y_1, \ldots, Y_n$ be a permutation of $X_1, \ldots, X_n$ such that*

$$Y_1 \leq Y_2 \leq \cdots \leq Y_n.$$

Then the ordered random variables $Y_1, \ldots, Y_n$ are called the order statistics *of the random sample $X_1, \ldots, X_n$. Here Y_k is called the* kth order statistic. *Because of continuity, the equality sign could be ignored.*

Remark. Although X_i's are iid random variables, the random variables Y_i's are neither independent nor identically distributed.

Thus, the minimum of X_i's is

$$Y_1 = \min(X_1, \ldots, X_n)$$

and the maximum is

$$Y_n = \max(X_1, \ldots, X_n).$$

The order statistics of the sample $X_1, X_2, \ldots, X_n$ can also be denoted by $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ where

$$X_{(1)} < X_{(2)} < \cdots < X_{(n)}.$$

Here $X_{(k)}$ is the kth order statistic and is equal to Y_k in Definition . One of the most commonly used order statistics is the median, the value in the middle position in the sorted order of the values.

Example

(i) The range $R = Y_n - Y_1$ is a function of order statistics.

(ii) The sample median M equals Y_{m+1} if $n = 2m+1$.

Hence, the sample median M is an order statistic, when n is odd. If n is even then the sample median can be obtained using the order statistic, $M = (1/2)\left[Y_{n/2} + Y_{(n/2)+1}\right]$.

The following result is useful in determining the distribution of functions of more than one order statistics.

Theorem *Let $X_1, \ldots, X_n$ be a random sample from a population with pdf $f(x)$. Then the joint pdf of order statistics $Y_1, \ldots, Y_n$ is*

$$f(y_1, \ldots, y_n) = \begin{cases} n! f(y_1) f(y_2) \ldots f(y_n), & \text{for } y_1 < \cdots < y_n \\ 0, & \text{otherwise.} \end{cases}$$

The pdf of the kth order statistic is given by the following theorem.

Theorem *The pdf of Y_k is*

$$f_k(y) = f_{Y_k}(y) = \frac{n!}{(k-1)!\,(n-k)!} f(y)\,(F(y))^{k-1}\,(1-F(y))^{n-k},$$

for $-\infty < y < \infty$, where $F(y) = P(X_i \le y)$ is the cdf of X_i.

In particular, the pdf of Y_1 is $f_1(y) = nf(y)[1-F(y)]^{n-1}$ and the pdf of Y_n is $f_n(y) = nf(y)[F(y)]^{n-1}$. In the following example, we will derive pdf for Y_n.

Example

Let $X_1, \ldots, X_n$ be a random sample from $U[0, 1]$. Find the pdf of the kth order statistic Y_k.

Solution

Since the pdf of X_i is $f(x) = 1, 0 \le x \le 1$, the cdf is $F(x) = x, 0 \le x \le 1$. Using Theorem the pdf of the kth order statistic Y_k reduces to

$$f_k(y) = \frac{n!}{(k-1)!\,(n-k)!} y^{k-1}\,(1-y)^{n-k}, \quad 0 \le y \le 1$$

which is a beta distribution with $\alpha = k$ and $\beta = n - k + 1$.

The next example gives the so-called extreme (i.e., largest) value distribution, which is the distribution of the order statistic Y_n.

Example

Find the distribution of the nth order statistic Y_n of the sample $X_1, \ldots, X_n$ from a population with pdf $f(x)$.

Solution

Let the cdf of Y_n be denoted by $F_n(y)$. Then

$$F_n(y) = P(Y_n \le y) = P\left(\max_{1 \le i \le n} X_i \le y\right)$$
$$= P(X_1 \le y, \ldots, X_n \le y) = [F(y)]^n \text{ (by independence).}$$

Hence, the pdf $f_n(y)$ of Y_n is

$$f_n(y) = \frac{d}{dy}[F(y)]^n = n[F(y)]^{n-1}\frac{d}{dy}F(y)$$
$$= n[F(y)]^{n-1} f(y).$$

In particular, if $X_1, \ldots, X_n$ is a random sample from $U[0, 1]$, then the cumulative extreme value distribution is given by

$$F_n(y) = \begin{cases} 0, & y < 0 \\ y^n, & 0 \le y \le 1 \\ 1, & y > 1. \end{cases}$$

Example

A string of 10 light bulbs is connected in series, which means that the entire string will not light up if any one of the light bulbs fails. Assume that the lifetimes of the bulbs, $\tau_1, \ldots, \tau_{10}$, are independent random variables that are exponentially distributed with mean 2. Find the distribution of the life length of this string of light bulbs.

Solution

Note that the pdf of τ_i *is* $f(t) = 2e^{-2t}$, $0 < t < \infty$, *and the cumulative distribution of* τ_i *is* $F_{\tau_i}(t) = 1 - e^{-2t}$. *Let* T *represent the lifetime of this string of light bulbs. Then,*

$$T = \min(\tau_1, \ldots, \tau_{10}).$$

Thus,

$$F_T(t) = 1 - [1 - F_{\tau_i}(t)]^{10}.$$

Hence, the density of T *is obtained by differentiating* $F_T(t)$ *with respect to* t, *that is,*

$$\begin{aligned} f_T(t) &= 10 f_{\tau_i}(t)[1 - F_{\tau_i}(t)]^9 \\ &= \begin{cases} 2(10)e^{-2t}(e^{-2t})^9 = 20e^{-20t}, & 0 < t < \infty \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

The joint pdf of the order statistics is given by the following result.

Theorem *Let $X_1, \ldots, X_n$ be a random sample with continuous probability density function $f(x)$ and a distribution function $F(x)$. Let $Y_1, \ldots, Y_n$ be the order statistics. Then for any $1 \leq i < k \leq n$ and $-\infty < x \leq y < \infty$, the joint pdf of Y_i and Y_k is given by*

$$f_{Y_i,Y_k}(x, y) = \frac{n!}{(i-1)!\,(k-i-1)!\,(n-k)!}\,[F(x)]^{i-1} \times [F(y) - F(x)]^{k-i-1}\,[1 - F(y)]^{n-k}\, f(x)\, f(y)$$

Example

Let $X_1, \ldots, X_n$ be a random sample from $U[0, 1]$. Find the joint pdf of Y_2 and Y_5.

Solution

Taking $i = 2$ and $k = 5$ in Theorem , we get the joint pdf of Y_2 and Y_5 as

$$f_{Y_2,Y_5}(x, y) = \frac{n!}{(2-1)!\,(5-2-1)!\,(n-5)!}\,[F(x)]^{2-1}\,[F(y) - F(x)]^{5-2-1} \times [1 - F(y)]^{n-5}\, f(x)\, f(y)$$

$$= \begin{cases} \frac{n!}{2(n-5)!}x\,(y-x)^2\,(1-y)^{n-5} & 0 < x \leq y < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور

فراس شاكر محمود

Point estimation

Definition: For a given positive integer n, $Y=(X_1,X_2,\ldots, X_n)$ will be called a best statistic for the parameter θ if Y is unbiased, E(Y)= θ, and if the variance of Y is less than or equal to the variance of every other unbiased statistic for θ.

Definition: There are many ways of defining a "best" statistic for a parameter . our Definition adopts the principles of biasedness and minimum variance as being reasonable. One of our purposes in adopting these principles is to motivate. In a somewhat natural way the study of an important class of statistics called "sufficient statistics estimator" stands for the value of that function; for example, $\bar{X}n = \sum Xi / n$ is an estimator of a mean μ function, and the word" estimate" stands for n and $\bar{X}n$ is an estimate of μ. Here T is $\bar{X}n$, and T(1 , ... , n) is the function defined by summing the arguments and then dividing by n. One of the basic problems is how to find an estimator of population parameter θ. There are several methods for finding an estimator of θ. Some of these methods are:

(1) Maximum Likelihood Method.
(2) Moment Method.
(3) Bayes Method.
(4) Least squares method.
(5) Minimum Chi – Squares Method.
(6) Minimum Distance Method.

Some properties of the estimator

To estimate a parameter of the population under study, we need to choose the appropriate statistic in the sample to estimate this parameter. Often the corresponding parameter in the sample is a better estimate, for example estimating the population mean μ through the sample mean m. The statistic used in the estimation is called the estimate.

Definition : The estimator is **unbiased**: We say of a statistic that it is an unbiased estimator of the population parameter if its mean or mathematical expectation is equal to the population parameter.

Example: We say about the sample mean, m, that it is an unbiased estimate of the population mean μ because $E(m) = \mu$. In contrast, we call the statistic S^2 in a return-sampling that it is a biased estimator of σ^2 because $E(S^2) = \sigma^2 (n-1)/n \neq \sigma^2$ while statistic $S'^2 = S^2 n / (n-1)$ is an unbiased estimator in a return preview

Definition : The estimator is **efficiency**: The efficiency of an estimator relates to the amount of variance of the sampling distribution of the statistic. If two (statistical) estimators have the same mean, we say that the estimator with the least disparate sampling distribution is the most efficient.

Example: For both the sampling distributions of the mean and the mean, the same mean is the population mean, but the mean m is considered a more efficient estimator of the population mean than the median because the variance of the sampling distribution of the averages $V(m) = \sigma^2/n$ is less than the variance of the sampling distribution for the median:

$$V(med) = \sigma^2\pi/2n = (\sigma^2/n)(3.14159/2) > \sigma^2/n$$

Obviously, using effective and unbiased capabilities is best, but other capabilities may be used to obtain them.

Definition : The estimator is **convergence** : We say an estimator is convergence if it refers to the estimated parameter value when the sample size tends to infinity.

Example: The sample mean is considered an convergence estimate of the population mean because:

$$E(m) = \mu \quad , \qquad V(m) = \frac{\sigma^2}{n} \xrightarrow[n \to \infty]{} 0.$$

Unbiased ness أ- عدم التحيز

Consistency ب- الاتساق

Mean Square Error (MSE) ج- متوسط مربعات الخطأ

Efficiency د- الكفاءة

Sufficiency ه- الكفاية

Completeness و- الكمال

1- Moment Method:

Let $X_1 X_2, \dots, X_n$ be a random sample from a population X with probability density function $f(x; \theta_1, \theta_2, \dots, \theta_m)$, where $\theta_1, \theta_2, \dots, \theta_m$ are m unknown parameters. Let

$$E(X^k) = \int_{-\infty}^{\infty} x^k f(x; \theta_1, \theta_2, \dots, \theta_m) dx$$

Be the k^{th} population moment about 0.

Further, let

$$M_k = \frac{1}{n} \sum_{i=1}^{n} X_i^k$$

Be the k^{th} sample moment about 0.

In moment method, we find the estimator for the parameters $\theta_1, \theta_2, \dots, \theta_m$ by equating the first m population moments (if they exist) to the first m sample moments, that is

$$E(X) = M_1$$

$$E(X^2) = M_2$$

$$E(X^3) = M_3$$

$$\vdots$$

$$E(X^m) = M_m$$

The moment method is one of the classical methods for estimating parameters and motivation comes from the fact that the sample moments are in some sense estimates for the population moments. The moment method was first discovered by British statistician Karl Pearson in 1902. Now we provide some examples to illustrate this method.

Example. Let $X \sim N(\mu, \sigma^2)$ and $X_1, X_2, \ldots, X_n$ be a random sample of size n from the population X. What are the estimators of the population parameters μ and σ^2 if we use the moment method?

Solution: Since the population is normal, that is

$$X \sim N(\mu, \sigma^2)$$

We know that

$$E(X) = \mu$$

$$E(X^2) = \sigma^2 + \mu^2$$

Hence

$$\mu = E(X)$$

$$= M_1$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$= \bar{X}.$$

Therefore, the estimator of the parameter μ is $\bar{X}$, that is

$$\hat{\mu} = \bar{X}.$$

Next, we find the estimator of σ^2 equating $E(X^2)$ to M_2. Note that

$$\sigma^2 = \sigma^2 + \mu^2 - \mu^2$$

$$= E(X^2) - \mu^2$$

$$= M_2 - \mu^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 .$$

The last line follows from the fact that

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2X_i\bar{X} + \bar{X}^2)$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \frac{1}{n}\sum_{i=1}^{n} 2X_i\bar{X} + \frac{1}{n}\sum_{i=1}^{n}\bar{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}\frac{1}{n}\sum_{i=1}^{n} X_i + \bar{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\bar{X}\bar{X} + \bar{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$$

Thus, the estimator of σ^2 is $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$, that is

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Example. Let $X_1, X_2, \dots, X_n$ be a random sample of size n from a population X whit probability density function

$$f(x;\theta) = \begin{cases} \theta x^{\theta-1} & if 0 < x < 1 \\ 0 & otherwise, \end{cases}$$

Where 0 < θ < ∞ is an unknown parameter. Using the method of moment find an estimator of θ ? If $x_1 = 0.2, x_2 = 0.6, x_3 = 0.5, x_4 = 0.3$ is a random sample of size 4, then what is the estimate of θ ?

Solution To find an estimator, we shall equate the population moment to the sample moment. The population moment E(X) is given by

$$E(X) = \int_0^1 x f(x;\theta) dx$$

$$= \int_0^1 x\theta x^{\theta-1} dx$$

$$= \theta \int_0^1 x^{\theta}\, dx$$

$$= \frac{\theta}{\theta+1}\left[x^{\theta+1}\right]_0^1$$

$$= \frac{\theta}{\theta+1}$$

We know that $M_1 = \bar{X}$. now setting M_1 equal to E(X) and solving for θ, we get

$$\bar{X} = \frac{\theta}{\theta+1}$$

That is

$$\theta = \frac{\bar{X}}{1-\bar{X}}$$

Where $\bar{X}$ is the sample mean. Thus, the statistic $\frac{\bar{X}}{1-\bar{X}}$ is an estimator of the parameter θ . Hence

$$\hat{\theta} = \frac{\bar{X}}{1-\bar{X}}$$

Since$x_1 = 0.2, x_2 = 0.6, x_3 = 0.5, x_4 = 0.3$, we have $\bar{X} = 0.4$ and

$$\hat{\theta} = \frac{0.4}{1 - 0.4} = \frac{2}{3}$$

Is an estimate of the θ .

Example. Let $X \sim poisson(\lambda)$ find Moment Estimate of λ ?

Solution: Since $X \sim poisson(\lambda)$

$$f(X) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{X!} & for\ x = 0,1,\dots \\ 0 & otherwise \end{cases}$$

$$E(x) = \lambda\,, Var(x) = \lambda\,, \bar{X} = \frac{\sum x_i}{n}$$

$$E(X) = \bar{X}$$

$$\lambda = \frac{\sum x_i}{n}$$

$$\hat{\lambda} = \bar{X}$$

Example: Let $\mathrm{X \sim Binomaill(20, p)}$ find Moment Estimate of p?

Solution: Since $\mathrm{X \sim Binomaill(20, p)}$

$$f(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & for\ x = 0,1,\dots,n \\ 0 & 0 \le p \le 1 \end{cases}$$

$$E(x) = np\,, Var(x) = npq$$

$$E(x) = 20p\,, \bar{X} = \frac{\sum x_i}{n}$$

$$E(x) = \bar{X} \rightarrow \frac{\left[20p = \frac{\sum x_i}{n}\right]}{20}$$

$$p = \frac{1}{20} \cdot \frac{\sum x_i}{n} \rightarrow p = \frac{\bar{X}}{20}$$

$\hat{p} = \frac{1}{20}\bar{X} = \frac{0.4}{20} = \frac{1}{50} = 0.02$

Example: Let $X_1, X_2, \dots, X_n$ a random variable sample from Uniform (0,θ) find the Moment Estimate of θ ?

Solution: Since U(0,θ)

$$f(x) = \begin{cases} \frac{1}{b-a} & for\ a \le x \le b \\ 0 & otherwise \end{cases}$$

$$E(x) = \frac{a+b}{2}$$

$$E(x) = \frac{0+\theta}{2}$$

$$E(x) = \bar{X}$$

$$\frac{\theta}{2} = \bar{X} \ \rightarrow \theta = 2\bar{X}$$

$$\hat{\theta} = 2\bar{X} = 2 \times 0.4 = 0.8$$

$$E(X^2) = \int X^2 \cdot f(x)dx \ \rightarrow \int_0^{\theta} X^2 \cdot \frac{1}{\theta} dx$$

$$E(X^2) = \left[\frac{1}{\theta} \cdot \frac{X^3}{3}\right]_0^{\theta} \rightarrow E(X^2) = \frac{\theta^3}{3\theta}$$

$$E(X^2) = \frac{\theta^2}{3}$$

$$E(X^2) = \frac{4\bar{X}^2}{3}$$

$$\frac{\theta^2}{3} = \frac{\sum X_i^2}{n} \ \rightarrow n\theta^2 = 3\sum X_i^2 \ \rightarrow \theta^2 = \frac{3\sum X_i^2}{n}$$

$$\hat{\theta} = \sqrt{\frac{3\sum X_i^2}{n}}$$

Example: Let $X_1, X_2, \dots, X_n$ be a random variable sample of size n from distribution, with p.d.f

$$f(x,\alpha,\theta) = \begin{cases} \frac{\alpha}{\theta^{\alpha}} X^{\alpha-1} & for\ 0 < x < \theta \\ 0 & otherwise \end{cases}$$

Where $\alpha > 0\ , \theta > 0$

Suppose α is known find the moment estimator of θ , $\hat{\theta}$ and unbiased estimator of θ ?

Solution:

$$E(X) = \int x \cdot f(x)dx \rightarrow E(X) = \int_0^{\theta} \frac{\alpha}{\theta^{\alpha}} x^{\alpha-1} \cdot xdx$$

$$E(X) = \int_0^{\theta} \frac{\alpha}{\theta^{\alpha}} \cdot x \cdot x^{-1} \cdot x^{\alpha} dx \rightarrow E(X) = \int_0^{\theta} \frac{\alpha}{\theta^{\alpha}} \cdot x^{\alpha} dx$$

$$E(X) = \left[\frac{\alpha}{\theta^{\alpha}} \cdot \frac{x^{\alpha+1}}{\alpha+1}\right]_0^{\theta} \rightarrow E(X) = \frac{\alpha(\theta)^{\alpha+1}}{\theta^{\alpha} \cdot (\alpha+1)} = \frac{\alpha\theta^{\alpha}\theta}{\theta^{\alpha}(\alpha+1)} = \frac{\alpha\theta}{\alpha+1}$$

$$E(X) = \bar{X}$$

$$\frac{\alpha\theta}{\alpha+1} = \bar{X} \quad \rightarrow \quad \frac{[\alpha\theta = (\alpha+1)\bar{X}]}{\alpha}$$

$$\hat{\theta} = \frac{(\alpha+1)\bar{X}}{\alpha} \qquad \rightarrow \qquad ,\bar{X} = \frac{\alpha\theta}{\alpha+1}$$

$$E(\hat{\theta}) = E\left[\frac{\alpha+1}{\alpha}\bar{X}\right]$$

$$\rightarrow \frac{\alpha+1}{\alpha} E(\bar{X})$$

$$\frac{\alpha+1}{\alpha}\left[\frac{\alpha\theta}{\alpha+1}\right]$$

$$E(\hat{\theta}) = \theta$$

$\hat{\theta}$ is unbiased estimator of θ .

2- Maximum Likelihood Estimator:

Let $L(\theta) = L(\theta; x_1,\ldots,x_n)$ be the likelihood function for the random variables $X_1, X_2, \ldots, X_n$. if θ (where $\theta = \vartheta(x_1, x_2, \ldots,x_n)$ is a function of the observations $x_1,\ldots,x_n$) is the Value of θ in Θ which maximum $L(\theta)$. Then $\Theta = \vartheta(X_1,X_2,\ldots,X_n)$ is the Maximum likelihood estimator of $\theta = \vartheta(x_1,\ldots,x_n)$ is the maximum likelihood estimate of θ for the example $x_1,\ldots,x_n$. The most likelihood important cases which we shall consider are those in which $X_1,X_2,\ldots,X_n$ is a random sample from some density $f(x;\theta)$, so that the likelihood function is

$$L(\theta) = f(x_1;\theta)\, f(x_2;\theta) \ldots f(x_n;\theta).$$

Many likelihood functions satisfy regularity conditions; so the maximum-likelihood estimator in the solution of the equation.

$$\frac{dL(\theta)}{d\theta} = 0$$

Also, $L(\theta)$ and $\log L(\theta)$ have their maxima at the same value of θ, and it is sometimes easier to find the maximum of the logarithm of the likelihood, if the likelihood function contains (k) parameters, that is

If $\quad L(\theta_1,\theta_2,\ldots,\theta_k) = \prod_{f=1}^{n} f(x_i; \theta_1, \theta_2, \ldots, \theta_k)$

Then the maximum-likelihood estimators of the parameters $\theta_1,\theta_2,\ldots, \theta_k$ are the random variables $\Theta_1 = \vartheta_1(X_1,\ldots,X_n), \ldots, \Theta_2 = \vartheta_2(X_1,\ldots,X_n), \ldots. \Theta_k = \vartheta_k(X_1,\ldots,X_n)$, where $\theta_1,\theta_2,\ldots,\theta_k$ are the values in Θ which maximize $L(\theta_1,\theta_2,\ldots,\theta_k)$.

If certain regularity conditions are satisfied, the point where the likelihood is a maximum is a solution of the (k) equation

$$\frac{dL(\theta_1, \ldots, \theta_k)}{d\theta_1} = 0$$

$$\frac{dL(\theta_2, \ldots, \theta_k)}{d\theta_2} = 0$$

$$\frac{dL(\theta_1, \ldots, \theta_k)}{d\theta_k} = 0$$

In this case it may also be easier to work with the logarithm of the likelihood,

We shall illustrate these definitions with some examples.

Example 1. Let $x_1, x_2,\dots, x_n$ a random variable sample ~ Geometric (p) find Maximum likelihood estimator of (p)

Solution:

Since $x_1,\dots, x_n$ ~ G(p)

$$f(x) = \begin{cases} P(1-P)^{x-1} & \text{For } x=1,2,\dots \quad 0 \le P \le 1 \\ 0 & \text{Otherwise} \end{cases}$$

$$f(x_1, \dots, x_n, P) = f(x_1, P) \, . f(x_2, P) \dots \dots . f(x_n, P)$$

$$f((x_1, \dots, x_n, P) = P(1-P)^{x_1-1} \, . P(1-P)^{x_2-1} \; \dots . P(1-P)^{x_n-1}$$

$$f(x_1, \dots, x_n, P) = \; P^n \, (1-P)^{\sum x_i - n}$$

$$\ln f(x_1, \dots, x_n, P) = \ln[P^n] + \ln(1-P)^{\sum x_i - n}$$

$$\ln f(x_1, \dots, x_n, P) = n \ln(P) + (\sum x_i - n) \ln(1-P)$$

$$\frac{d \ln f(x_1,\dots,x_n,P)}{dP} = n . \frac{1}{P} + (\sum x_i - n) . \frac{-1}{1-P} = 0$$

$$\left[\frac{n}{P} - (\sum_{i=1}^{n} x_i - n) \frac{1}{1-P}\right] = 0 \qquad * \; \frac{1-P}{n}$$

$$\frac{1-P}{P} = \frac{\sum x_i - n}{n}$$

$$\frac{1}{P} - \frac{P}{P} = \frac{\sum_{i=1}^{n} x_i}{n} - \frac{n}{n} ==> \frac{1}{P} - 1 = \frac{\sum x_i}{n} - 1$$

$$\frac{1}{P} = \frac{\sum x_i}{n} ==> P \sum_{i=1}^{n} x_i = n$$

$$P = \frac{n}{\sum_{i=1}^{n} x_i} ==> P = \frac{1}{\bar{x}}$$

$$\hat{P} = \frac{1}{\bar{x}}$$

Example Let $x_1, \dots, x_n$ ~ Poisson (λ) find Maximum-likelihood estimator of λ?

Solution: Since $x_1, \dots, x_n$ ~ Poisson (λ)

$$f(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & \text{For } x=0,1,\dots \\ 0 & \text{Otherwise} \end{cases}$$

$$f(x_1, \dots, x_n, \lambda) = f(x_1, \lambda) \,.\, f((x_2, \lambda) \dots \dots f((x_n, \lambda)$$

$$f(x_1, \dots, x_n, \lambda) = \frac{\lambda^{x_1} e^{-\lambda}}{x_1!} \,.\, \frac{\lambda^{x_2} e^{-\lambda}}{x_2!} \dots \dots . . \frac{\lambda^{x_n} e^{-\lambda}}{x_n!}$$

$$g(x_n, \lambda) = \frac{e^{-n\lambda} \lambda^{\sum_{1=1}^{n} x_i}}{\sum_{i=1}^{n} x_i \,!}$$

$$\ln g(x_n, \lambda) = \ln \left[\frac{e^{-n\lambda} \lambda^{\sum x_i}}{\sum x_i !} \right]$$

$$\ln g\,(x_n, \lambda) = \ln e^{-n\lambda} + \ln \lambda^{\sum x_i} - \ln \sum_{i=1}^{n} x_i \,!$$

$$\ln g\,(x_n, \lambda) = -n\lambda + \sum_{i=1}^{n} x_i \ln \lambda - 0$$

$$\ln g(x_n, \lambda) = -n\lambda + \sum_{i=1}^{n} x_i \ln \lambda$$

$$\frac{d \ln g(x_n, \lambda)}{\lambda} = 0$$

$$\frac{d \ln g(x_n, \lambda)}{\lambda} = -n + \sum_{i=1}^{n} x_i \ . \frac{1}{\lambda}$$

$$-n + \sum_{i=1}^{n} x_i \ . \frac{1}{\lambda} = 0$$

$$\sum_{i=1}^{n} x_i \frac{1}{\lambda} = n ==> \left[n\lambda = \sum_{i=1}^{n} x_i \right] \div n$$

$$\lambda = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\hat{\lambda} = \bar{x}$$

<u>Example</u>: Let X ~ Bernoulli Parameters (P) find Maximum likelihood estimator of P ?

<u>Solution</u>: Since X ~ Ber (P)

$$f(x) = \begin{cases} P^x (1-P)^{1-x} & \text{For x=0,1} \\ 0 & \text{Otherwise} \end{cases}$$

$$f(x_1, \ldots, x_n, P) = f(x_1, P). f(x_2, P) \ldots \ldots f(x_n, P)$$

$$f(x_1, \ldots., x_n, P) = P^{x_1}(1-P)^{1-x_1}. P^{x^2}(1-P)^{1-x_2} \ldots.. P^{x_n}(1-P)^{1-x_n}$$

$$g(x_n, P) = P^{\sum_{i=1}^{n} x_i}(1-P)^{n-\sum_{i=1}^{n} x_i}$$

$$\text{Ln}\, g(x_n, P) = \ln\left[P^{\sum x_i}\right] + \ln\left[(1-P)^{n-\sum x_i}\right]$$

$$\ln g(x_n, P) = \sum_{i=1}^{n} x_i \ln(P) + (n - \sum_{i=1}^{n} x_i) \ln(1-P)$$

$$\frac{d \ln g(x_{n1}P)}{dP} = \sum_{i=1}^{n} x_i \frac{1}{P} + (n - \sum_{i=1}^{n} x_i) \frac{-1}{1-P} = \sum_{i=1}^{n} x_i \frac{1}{P} - (n - \sum_{i=1}^{n} x_i) \frac{1}{1-P}$$

$$\frac{d \ln g(x_n, P)}{dP} = \sum_{i=1}^{n} x_i \frac{1}{P} - (n - \sum_{i=1}^{n} x_i) \frac{1}{1-P}$$

$$\frac{d \ln g(x_n, P)}{dP} = 0$$

$$\sum_{i=1}^{n} x_i \frac{1}{P} - (n - \sum_{i=1}^{n} x_i) \frac{1}{1-P} = 0$$

$$\left[\sum x_i \frac{1}{P} = (n - \sum_{i=1}^{n} x_i) \frac{1}{1-P} \right] * \frac{1-P}{\sum_{i=1}^{n} x_i}$$

$$\frac{1-P}{P} = \frac{n - \sum x_i}{\sum x_i} - 1$$

$$\frac{1-P}{P} = \frac{P}{P} = \frac{n}{\sum_{i=1}^{n} x_i} - \frac{\sum x_i}{\sum x_i} ==> \frac{1}{P} - 1 = \frac{n}{\sum_{i=1}^{n} x_i} - 1$$

$$\frac{1}{P} = \frac{n}{\sum_{i=1}^{n} x_i} ==> \left[nP = \sum_{i=1}^{n} x_i \right] \div n$$

$$P = \frac{\sum_{i=1}^{n} x_i}{n} ==> P = \bar{\chi}$$

$$\hat{P} = \bar{\chi}$$

Example: Let $x_1 \dots x_n$ ~ N(μ, σ^2) find Maximum-likelihood estimator of μ & σ^2?

Solution:

Since $x_1, \ldots., x_n \sim$ N(μ, σ^2) then $f(x) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} & \text{for } -\infty \leq x \leq \infty \\ 0 & \text{otherwise} \end{cases}$

$$f(x_1, \ldots., x_n, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1(x_1-\mu)^2}{2\sigma^2}} . \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1(x_2-\mu)^2}{2\sigma^2}} \ldots \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1(x_n-\mu)^2}{2\sigma^2}}$$

$$g(x_1, \ldots. x_n, \mu, \sigma^2) = \frac{1}{\sigma^n \sqrt{(2\pi)^n}} e^{\frac{-1(\sum x_i-\mu)^2}{2\sigma^2}}$$

$$g(x_n, \mu, \sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} e^{\frac{-(\sum x_i-\mu)^2}{2}}$$

$$\ln g(x_n, \mu, \sigma^2) = (2\pi)^{\frac{-n}{2}} + \ln(\sigma^2)^{\frac{-n}{2}} + \ln e^{\frac{-(\sum x_i-\mu)^2}{2\sigma}}$$

$$\ln g(x_n, \mu, \sigma^2) = \frac{-n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{(\sum x_i - \mu)^2}{2\sigma^2}$$

$$\frac{d \ln g(x_n, \mu, \sigma^2)}{d\mu} = 0 - 0 - \frac{1}{2\sigma^2}(-2)(\sum x_i - \mu)$$

$$\frac{d \ln g(x_n \mu)}{d\mu} = \frac{+2(\sum x_i - \mu)}{2\sigma^2}$$

$$\frac{(\sum x_i - \mu)}{\sigma^2} = 0 ==> (\sum x_i - \mu) = 0$$

$$\sum x_i = n.\mu ==> \mu = \frac{\sum x_i}{n} \qquad \mu = \bar{x} \rightarrow \hat{\mu} = \bar{x}$$

$$\ln g(x_n, \bar{x}, \sigma^2) = \frac{-n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{(\sum x_i - \bar{x})^2}{2\sigma^2}$$

$$\frac{d \ln g(x_n, \bar{x}, \sigma^2)}{d\sigma^2} = 0 - \frac{n}{2} . \frac{1}{\sigma^2} - \frac{0 - 2(\sum x_i - \bar{x})^2}{4\sigma^4}$$

$$\frac{d \ln g(x_n, \bar{x}, \sigma^2)}{d\sigma^2} = \frac{-n}{2\sigma^2} + \frac{(\sum x_i - \bar{x})^2}{2\sigma^4}$$

$$\frac{-n}{2\sigma^2} + \frac{(\sum x_i - \bar{x})^2}{2\sigma^4}) = 0$$

$$\frac{\left(\sum x_i - \bar{x}\right)^2}{2\sigma^4} = \frac{n}{2\sigma^2} ==> \left[2\sigma^2 \sum_{i=1}^{n} (x_i - \bar{x})^2 = 2n\sigma^4\right] \div 2\sigma^2 n$$

$$\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} = \sigma^2 ==> \hat{\sigma}^2 = S^2$$

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور

فراس شاكر محمود

The Unbiased Estimator

Let $X_1, X_2, \dots \dots, X_n$ be a random sample of size n from apopulation with probability density function $f(x{:}\,\theta)$. An estimator $\hat{\theta}$ of θ is a function of the random variables $X_1, X_2, \dots \dots, X_n$ which is free of the parameter θ .

An estimate is a realized value of an estimator that is obtained when a sample is actually taken .

Definition: An estimator $\hat{\theta}$ of θ is said to be an unbiased estimator of θ if and only if

$$\boldsymbol{E(\hat{\theta}) = \theta}$$

If $\hat{\theta}$ is not unbiased , then it is called a biased estimator of θ .

An estimator of a parameter may not equal to the actual value of the parameter for every realization of the sample $X_1, X_2, \dots \dots, X_n$, but if it is unbiased then on an average it will equal to the parameter .

Example: Let $X_1, X_2, \dots \dots, X_n$ be a random sample from a normal population with mean μ and variance $\sigma^2 > 0$. Is the sample mean $\bar{X}$ an unbiased estimator of the parameter μ ?

Solution: Since , each $X_i \sim N(\mu, \sigma^2)$, we have $\bar{X} \sim N(\mu\,, \frac{\sigma^2}{n})$.

That is , the sample mean is normal with mean μ and variance $\frac{\sigma^2}{n}$.

Thus $E(\bar{X}) = \mu$. Therefore, the sample mean $\bar{X}$ is an unbiased estimator of μ

Example: Let $X_1, X_2, \dots \dots, X_n$ be a random sample from a population with mean μ and variance $\sigma^2 > 0$,

Is the sample variance S^2 an unbiased estimator of the population variance σ^2 ?

Solution: Note that the distribution of the population is not given . However , we are given $\mathrm{E}(\bar{\mathrm{X}}) = \mu$ and $\mathrm{E}[(\mathrm{X} - \mu)^2] = \sigma^2$.

In order to find $E(S^2)$, we need find $E(\bar{X})$ and $E(\bar{X}^2)$. Thus we proceed to find these two expected values .

$$E(\bar{X}) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$

Similarly:

$$Var(\bar{X}) = Var\left(\frac{X_1 + X_2 + \cdots X_n}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}$$

Therefore

$$E(\bar{X}^2) = Var(\bar{X}) + E(\bar{X})^2 = \frac{\sigma^2}{n} + \mu^2$$

Consider

$$E(S^2) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right]$$

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n}(X_i^2 - 2\bar{X}X_i + \bar{X}^2)\right]$$

$$= \frac{1}{n-1} \mathrm{E}\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]$$

$$= \frac{1}{n-1}\left\{\sum_{i=1}^{n} E[X_i^2] - E[n\bar{X}^2]\right\}$$

$$= \frac{1}{n-1}\left[n(\sigma^2 + \mu^2) - n(\mu^2 + \frac{\sigma^2}{n})\right]$$

$$= \frac{1}{n-1}[(n-1)\sigma^2]$$

$$E(S^2) = E(\hat{\sigma}^2) = \sigma^2 \quad .$$

Therefore , the sample variance S^2 is an unbiased estimator of the population variance σ^2 .

Example: If $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ and let S_1^2, S_2^2 are estimators of σ^2 , Show that S_1^2 is unbiased estimators of σ^2 and S_2^2 is biased estimator of σ^2. Such that :

$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \text{ and } S_2^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Solation:

$$Z = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \rightarrow E(Z) = (n-1)$$

$$\therefore Z_1 = \frac{(n-1)S_1^2}{\sigma^2} \sim \chi^2(n-1) \rightarrow E(Z_1) = (n-1) \text{ \_\_\_\_\_(1)}$$

$$E(Z_1) = E\left(\frac{(n-1)S_1^2}{\sigma^2}\right) = \frac{(n-1)}{\sigma^2}E(S_1^2) \text{ \_\_\_\_\_\_\_(2)}$$

From (1) and (2)

$$\frac{(n-1)}{\sigma^2}E(S_1^2) = (n-1) \rightarrow E(S_1^2) = \sigma^2$$

$\therefore S_1^2$ is an unbiased estimator of σ^2 .

$$Z_2 = \frac{nS_2^2}{\sigma^2} \sim \mathrm{X}^2(\mathrm{n}-1) \rightarrow \mathrm{E}(Z_2) = (n-1) \text{ \_\_\_\_\_(3)}$$

$$E(Z_2) = E\left(\frac{nS_2^2}{\sigma^2}\right) = \frac{n}{\sigma^2}E(S_2^2) \text{ \_\_\_\_\_\_(4)}$$

From (3) and (4) $\frac{n}{\sigma^2}E(S_2^2) = n-1 \rightarrow E(S_2^2) = \frac{(n-1)}{n}\sigma^2$

S_2^2 is a biased estimator of σ^2 .

Example: Let $X_1, X_2, \dots, X_n$ be a random sample from a Bernoulli population with parameter p , show that $\bar{X}_n$ is an unbiased estimator.

Solation:

$$E(\bar{X}_n) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}(E(X_i)) = \frac{1}{n}\left(\sum_{i=1}^{n} p\right)$$

$$= \frac{1}{n} np = p$$

Then $E(\hat{p}) = E(\bar{\mathrm{X}}_{\mathrm{n}}) = p$ is an unbiased estimator for p .

Example*:* Let X_1, X_2 and X_3 be a sample of size $n = 3$ from a distribution with unknown mean $-\infty < \mu < \infty$, and the variance σ^2 is a known positive number. Show that both $\hat{\theta}_1 = \bar{X}$ and $\hat{\theta}_2 = \frac{1}{8}(2X_1 + \mathrm{X}_2 + 5\mathrm{X}_3)$ are unbiased estimator for μ . Compare the variance of $\theta_1^{\wedge}$ and $\theta_2^{\wedge}$.

Solution :

$$E(\hat{\theta}_1) = E(\bar{X}) = E\left(\frac{1}{3}\sum_{i=1}^{3} X_i\right) = \frac{1}{3}\, 3\mu = \mu$$

$$E(\hat{\theta}_2) = \frac{1}{8}E(2X_1 + X_2 + 5X_3) = \frac{1}{8}[2E(X_1) + E(X_2) + 5E(X_3)]$$

$$= \frac{1}{8}(2\mu + \mu + 5\mu) = \frac{1}{8}(8\mu) = \mu$$

$\therefore\ \hat{\theta}_1 , \hat{\theta}_2$ are unbiased estimators .

$$Var(\hat{\theta}_1) = V\left(\frac{1}{3}\sum_{i=1}^{3} X_i\right) = \frac{1}{9}[V(X_1) + V(X_2) + V(X_3)]$$

$$= \frac{1}{9}[\sigma^2 + \sigma^2 + \sigma^2] = \frac{1}{9}\, 3\sigma^2 = \frac{1}{3}\sigma^2$$

$$Var(\hat{\theta}_2) = V\left[\frac{1}{8}(2X_1 + X_2 + 5X_3)\right]$$

$$= \frac{1}{64}[V(2X_1 + X_2 + 5X_3)]$$

$$= \frac{1}{64}[4V(X_1) + V(X_2) + 25V(X_3)]$$

$$= \frac{1}{64}(4\sigma^2 + \sigma^2 + 25\sigma^2) = \frac{1}{64}(30\sigma^2)$$

$$= \frac{15}{32}\sigma^2$$

Factorization (jointly sufficient statistics)

Theorem : Let $X_1, X_2, \dots\dots, X_n$ be a random sample of size n from the density $f(.;\theta)$, where the parameter θ may be a vector . A set of statistics

$$S_1 = \sigma_1(X_1, X_2, \dots\dots, X_n), \dots\dots\dots\dots S_r = \sigma_r(X_1, X_2, \dots\dots, X_n).$$

Is jointly sufficient if and only if the joint density of $X_1, X_2, \dots\dots, X_n$ can be factored as $f_{X_1,\dots\dots X_n}(X_1, X_2, \dots\dots, X_n ;\theta)$

$$= g(\sigma_1(X_1, X_2, \dots\dots, X_n), \dots\dots, \sigma_r(X_1, X_2, \dots\dots, X_n);\theta)$$

$$= g(S_1, \dots\dots, S_r ;\ \theta)\ h(X_1, X_2, \dots\dots, X_n),$$

where the function $h(X_1, X_2, \dots\dots, X_n)$ is nonnegative and does not involve the parameter θ and the function $g(S_1, \dots\dots, S_r ;\ \theta)$ is nonnegative and depends on $(X_1, X_2, \dots\dots, X_n)$ only through the functions $\sigma_1(., \dots\dots, .), \dots., \sigma_r(., \dots., .)$.

Note that , according to Theorem . There are many possible sets of sufficient statistics. The above two theorems give us a relatively easy method for judging whether a certain statistic is sufficient or a set of statistics is jointly sufficient .

However , the method is not the complete answer since a particular statistic may be sufficient yet the user may not be clever enough to factor the joint density .

The theorems may also be useful in discovering sufficient statistics . Actually , the result of either of the above factorization theorems is intuitively evident if one notes the following:

1- If the joint density factors as indicated , then the likelihood function is proportional to $g(S_1, \ldots\ldots, S_r;\ \theta)$, which depends on the observations $X_1, \ldots\ldots, X_n$ only through $\sigma_1, \ldots\ldots, \sigma_r$ [the likelihood function is viewed as a function of θ, so $h(X_1, X_2, \ldots\ldots, X_n\ )$ is just a proportionality constant] , which means that the information about θ that the likelihood function contains is embodied in the statistics

$\sigma_1(., \ldots\ldots, .), \ldots., \sigma_r(., \ldots., .)$.

Example: $\sum_{i=1}^{n} X_i$ is sufficlent to μ , $X_i \sim (\mu, \sigma^2)$ by using faclorization theorm .

Solation:

$$f[X_1, \ldots\ldots, X_n, \mu] = f(X_1, \mu).f(X_2, \mu). \ \ldots. \ .f(X_n, \mu)$$

Since $X_i \sim (\mu, \sigma^2)$

$$\therefore f(X) = \left\{\frac{1}{\sigma\sqrt{2\pi}}\ e^{\frac{-1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}\right\} \quad \text{for} \quad -\infty < X < \infty$$

$$f[X_1, \ldots\ldots, X_n, \mu] = f(X_1, \mu).f(X_2, \mu). \ \ldots. \ .f(X_n, \mu)$$

$$\sum \frac{1}{\sigma\sqrt{2\pi}}\ e^{\frac{-1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} . \frac{1}{\sigma\sqrt{2\pi}}\ e^{\frac{-1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} \ .\ldots. \ .\frac{1}{\sigma\sqrt{2\pi}}\ e^{\frac{-1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n . e^{\frac{-1(\sum X_i - \mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n . e^{\frac{[(\sum X_i)^2 - 2\mu \sum X_i + \mu^2]}{2\sigma^2}}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n . \ e^{\frac{-(\sum X_i)^2}{2\sigma^2}} \ . e^{\frac{-(-2\mu \sum X_i + \mu)}{2\sigma^2}}$$

$$h(X) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n . \ e^{\frac{-(\sum X_i)^2}{2\sigma^2}} \quad ,$$

$$g(t(X),\theta) = e^{\frac{-(-2\mu\sum X_i+\mu)}{2\sigma^2}}$$

$\therefore \sum X_i$ is sufficient statistic to μ .

<u>Example:</u> $\sum_i^n X_i$ is sufficient statistic to 1 $X \sim (1, \sigma^2)$ by using faclorization theorem

<u>Solation:</u>

$$X_i \sim N(1, \sigma^2)$$

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} . e^{\frac{-1(X-1)^2}{2\sigma^2}}$$

Since X_i is i.i.d

$$f(X_1, \ldots, X_n, 1) = f(X_1, 1).f(X_2, 1). \ \ldots \ .f(X_n, 1)$$

$$f(X_1, \ldots, X_n, 1) = \frac{1}{\sigma\sqrt{2\pi}} . e^{\frac{-1(X-1)^2}{2\sigma^2}} \ldots \frac{1}{\sigma\sqrt{2\pi}} . e^{\frac{-1(X-1)^2}{2\sigma^2}} .$$

$$f(X_1, \ldots, X_n, 1) = \left[\frac{1}{\sigma\sqrt{2\pi}}\right]^n . e^{\frac{-1(\sum X_i-1)^2}{2\sigma^2}}$$

$$f(X_1, \ldots, X_n, 1) = \left[\frac{1}{\sigma\sqrt{2\pi}}\right]^n . e^{\frac{-1[(\sum X_i)^2-2\sum X_i+1]}{2\sigma^2}}$$

$$f(X_1, \ldots, X_n, 1) = \left[\frac{1}{\sigma\sqrt{2\pi}}\right]^n . e^{\frac{-1(\sum X_i)^2}{2\sigma^2}} . e^{\frac{-(-2\sum X_i+1)}{2\sigma^2}}$$

$$h(X) = \left[\frac{1}{\sigma\sqrt{2\pi}}\right]^n . e^{\frac{-1(\sum X_i)^2}{2\sigma^2}} \quad ,$$

$$g(t(x),\theta) = . e^{\frac{-(-2\sum X_i+1)}{2\sigma^2}}$$

$\therefore \sum_i^n X_i$ is sufficient statistic to 1 .

<u>Example</u>: $\boldsymbol{\sum_{i=1}^n X_i}$ is sufficient statistic to γ ,

$X_i \sim pio(\gamma)$ by using faclorization .

Solation:

Since $X_i \sim pio(\gamma)$

$$f(X) = \left\{ \frac{\gamma^X e^{-\gamma}}{X!} \quad for\ X = 0,1,,,,\infty \right\}$$

Since X_i is $i.i.d$

$$f[X_1, \ldots\ldots, X_n, \gamma] = f(X_1, \gamma).f(X_2, \gamma).\ \ldots.\ .f(X_n, \gamma)$$

$$f[X_1, \ldots\ldots, X_n, \gamma] = \frac{\gamma^{X_1} e^{-\gamma}}{X_1!} \ .\ \frac{\gamma^{X_2} e^{-\gamma}}{X_2!} \ldots\ldots\ldots$$

$$f[X_1, \ldots\ldots, X_n, \gamma] = \frac{\gamma^{\sum X_i}\ e^{-\gamma}}{\sum_{i=1}^{n} X_i}$$

$$f[X_1, \ldots\ldots, X_n, \gamma] = \frac{1}{\sum_{i=1}^{n} X_i} \ .\left(\gamma^{\sum X_i}\ e^{-\gamma}\right)$$

$$h(X) = \frac{1}{\sum_{i=1}^{n} X_i} \quad ,$$

$$g(t(X), \theta) = \left(\gamma^{\sum X_i}\ e^{-\gamma}\right)$$

$\therefore$ $\sum_{i=1}^{n} X_i$ is sufficient statistic to γ.

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

Department of Mathematics

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد

الدكتور فراس شاكر محمود

Mean square error متوسط مربعات الخطأ

Definition:

The mean square error of the estimator θ, denoted by MSE (θ) is defined as

MSE $(\hat{\theta})$ = $E(\hat{\theta}-\theta)^2$ = Var $(\hat{\theta})$ + $[E(\hat{\theta}) - \theta]^2$

$$\begin{aligned} \text{MSE}(\hat{\theta}) &= \text{E}(\hat{\theta}-\theta)^2 \\ &= \text{E}\{ [\ \hat{\theta}-\text{E}(\hat{\theta})\ ] \ + [\ \text{E}(\hat{\theta})-\theta\ ]\}^2 \\ &= \text{E}[\hat{\theta}-\text{E}(\hat{\theta})]^2 + \text{E}[\text{E}(\hat{\theta})-\theta]^2 \\ &\qquad + \underbrace{2\,\text{E}\left\{[\hat{\theta}-\text{E}(\hat{\theta})][\text{E}(\hat{\theta})-\theta]\right\}}_{=0} \\ &= \text{E}[\hat{\theta}-\text{E}(\hat{\theta})]^2 + \text{E}[\text{E}(\hat{\theta})-\theta]^2 \\ &= \text{Var}(\hat{\theta}) + \underbrace{[\text{E}(\hat{\theta})-\theta]^2}_{\text{Bias}(\hat{\theta})^2} \end{aligned}$$

Definition:

The unbiased estimator $\hat{\theta}$ that minimizes the mean square error is called the minimum variance unbiased estimator (MVUE) of θ.

Example:

Let X_1, X_2, X_3 be a sample of size n= 3 from a distribution with unknown mean μ, $-\infty < \mu < \infty$, where the variance σ^2 is a known positive number.Show that both $\widehat{\theta_1} = \bar{X}$ and $\widehat{\theta_2} = [(2X_1 + X_2 + 5X_3)/8]$ are unbiased estimators for μ. Compare the variances of $\widehat{\theta_1}$, and $\widehat{\theta_2}$.

Solution:

We have E$(\widehat{\theta_1})$ = E $(\bar{X}) = \frac{1}{3}\, 3\,\mu = \mu$. And E$(\widehat{\theta_2})$ = E $[(2X_1 + X_2 + 5X_3)/8]$

$= \frac{1}{8}\,[(2E(X_1) + E(X_2) + 5E(X_3)] = \frac{1}{8}\,[\,2\mu + \mu + 5\,\mu\,) = \mu$

Hence, both $\widehat{\theta_1}$, and $\widehat{\theta_2}$, are unbiased estimators. However، Var $(\widehat{\theta_1})$ = var $(\bar{X})$ = $\frac{1}{3}\,\sigma^2$. Whereas Var $(\widehat{\theta_2})$ = var [($2X_1$ + X_2 + $5X_3$)/8]

= $\frac{1}{64}$ [4 var (X_1) + var (X_2) + 25 var (X_3) = $\frac{1}{64}$ 30 σ^2

Because var $(\widehat{\theta_1})$ < var $(\widehat{\theta_2})$, we see that $\bar{X}$ is a better unbiased estimator in the sense that the variance of $\bar{X}$ is smaller.

ملاحظات

- ❖ في حالة التقدير غير متحيز يكون $[E(\hat{\theta})-\theta]=0$

وبالتالي فإن $MSE(\hat{\theta})=Var(\hat{\theta})$

- ❖ نفرض أن $Bias(\hat{\theta})=B(\hat{\theta})=E(\hat{\theta})-\theta$ بشكل عام (متحيز أو غير متحيز)
- ❖ $MSE(\hat{\theta})=Var(\hat{\theta})+B(\hat{\theta})^2$
- ❖ الكفاءة بين مقدرين $e(\theta_1,\theta_2).=\frac{MSE(\widehat{\theta_2})}{MSE(\widehat{\theta_1})}<1$ فيكون المقدر الثاني اكفأ من المقدر الاول.

Example:

If $x_1,\ldots,x_n\sim N(M,\sigma^2)$ consider the two estimators of σ^2, $\widehat{\theta_1}={s_1}^2=\frac{1}{(n-1)}\sum(x_i-\bar{x})^2$, $\widehat{\theta_2}={s_2}^2=\frac{1}{n}\sum(x_i-\bar{x})^2$. Find the $e(\theta_1,\theta_2)$.

Solution :

$$E({s_1}^2)=\sigma^2\ \Rightarrow MSE({s_1}^2)=var\ ({s_1}^2)$$

$$var\left(\frac{(n-1){s_1}^2}{\sigma^2}\right)=2(n-1)\ \Rightarrow\ \frac{(n-1)^2}{\sigma^4}\ v({s_1}^2)=2(n-1)$$

$$\Rightarrow v({s_1}^2)=\frac{2\sigma^4}{(n-1)}=MSE\ ({s_1}^2)$$

للتوضيح

$$v({s_2}^2)=\frac{2(n-1)}{n^2}\sigma^4\qquad,\ \ E({s_2}^2)=\frac{(n-1)}{n}\sigma^2$$

$$B({s_2}^2)=\ E({s_2}^2)-\sigma^2=\frac{(n-1)}{n}\sigma^2-\sigma^2=\sigma^2-\frac{1}{n}\sigma^2-\sigma^2=-\frac{1}{n}\sigma^2$$

$$MSE(s_2{}^2) = v(s_2{}^2) + \left(B(s_2{}^2)\right)^2 = \frac{(2n-2)\sigma^4}{n^2} + \frac{1}{n^2}\sigma^4 = \frac{(2n-2+1)\sigma^4}{n^2}$$
$$= \frac{(2n-1)\sigma^4}{n^2}$$

$$e = \frac{MSE(s_2{}^2)}{MSE(s_1{}^2)} = \frac{\frac{(2n-1)}{n^2}\sigma^4}{\frac{2}{(n-1)}\sigma^4} = \frac{(2n-1)(n-1)}{2n^2} < 1$$

$s_2{}^2$ is relatively more efficient than $s_1{}^2$.

Definition:

اذا كان $\hat{\theta}$ تقدير غير متحيز لـ θ وكان $v(\hat{\theta}) = \frac{1}{nE\left[-\frac{\partial^2 \ln(f)}{\partial \theta^2}\right]}$، يكون $\hat{\theta}$ التقدير الغير متحيز ذو أقل تباين المنتظم ويرمز له (Uniformly Minimum Variance Unbiased Estimator) (UMVUE)

Example: let $x_1, \dots \dots \dots, x_n \sim N(M, \sigma^2)$ show that $\bar{x}$ is an efficient estimator.

Solution :

$$f_{(x)} = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$$\ln(f) = \ln\left(\frac{1}{\sqrt{2\pi}\ \sigma}\right) - \frac{1}{2\sigma^2}(x-\mu)^2$$

$$\frac{\partial \ln(f)}{\partial \mu} = 0 - \frac{1}{2\sigma^2}\ 2(x-\mu)\ (-1) = \frac{(x-\mu)}{\sigma^2} = \frac{x}{\sigma^2} - \frac{\mu}{\sigma^2}$$

$$\frac{\partial^2 \ln(f)}{\partial \mu^2} = -\frac{1}{\sigma^2}$$

$$\frac{1}{nE\left[-\frac{\partial^2 \ln(f)}{\partial \mu^2}\right]} = \frac{1}{nE\left[-\frac{1}{\sigma^2}\right]} = \frac{1}{n\frac{1}{\sigma^2}} = \frac{\sigma^2}{n} = v(\bar{x})$$

$\bar{x}$ is an efficient estimator of μ

$\bar{x}$ is an UMVUE of μ

❖ اذا كان التقديرين غير متحيز يكون $e(\theta_1, \theta_2) = \frac{v(\theta_2)}{v(\theta_1)}$

❖ اذا كان التقديرين بشكل عام سواء (متحيز أو غير متحيز) نستخدم القانون

$e(\theta_1 , \theta_2) = \frac{MSE\ (\theta_2)}{MSE\ (\theta_1)}$

❖ حل الأمثلة لتقدير المتسق بالطريقة الثانية.

Example: let $x_1, \dots, x_n \sim Poi(\lambda)$ show that $\bar{x}$ is an consistent estimator of the (λ) .

Solution :

1) $E(\bar{x}) = \frac{1}{n}\big(E(x_1) + \cdots + E(x_n)\big) = \frac{1}{n}(\lambda + \cdots + \lambda)$

$$\text{n} - \text{times} \qquad = \frac{1}{n}(n\lambda) = \lambda$$

2) $v(\bar{x}) = \frac{\lambda}{n}$

$$\lim_{n\to\infty} v(\bar{x}) = \lim_{n\to\infty} \frac{\lambda}{n} = 0$$

$\bar{x}$ is a consistent estimator of λ.

Example: let $x_1, \dots \dots \dots, x_n \sim N(\mu, \sigma^2)$

a) show that the sample variance s^2 is a consistent estimator for σ^2.
b) Show that the maximum likelihood estimator (MLE) for μ & σ^2 are consistent estimator for μ & σ^2

Solution :

a)

1) $E(s^2) = \sigma^2$

2) $v(s^2) = \frac{2\sigma^4}{n-1}$

$$\lim_{n\to\infty} v(s^2) = \lim_{n\to\infty} \frac{2\sigma^4}{n-1} = 0$$

s^2 is consistent estimator of σ^2

b)

MLE $\hat{\mu}$=$\bar{X}$ & MLE σ^2= $\frac{1}{n}\sum(X_i - \bar{X})^2$

1) $E(\bar{X}) = \mu$
2) $V(\bar{X}) = \frac{\sigma^2}{n} \quad \Rightarrow \quad \lim_{n\to\infty} v(\bar{X}) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$

$\bar{X}$ is consistent estimator of μ

MLE $\hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \bar{X})^2$

$E(\hat{\sigma}^2) = E(\frac{1}{n}\sum(X_i - \bar{X})^2) = \frac{(n-1)}{n}\left[E\frac{\sum(X_i-\bar{X})^2}{(n-1)}\right] = \frac{(n-1)}{n}\sigma^2$

∴ σ^2 is biased

$$Z = \frac{(n-1)}{\sigma^2} S^2 \sim X^2(n-1)$$

$$E(Z) = (n-1)$$

$$V(Z) = 2(n-1)$$

$$B(\hat{\sigma}^2) = E(\hat{\sigma}^2) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = \left(1 - \frac{1}{n}\right)\sigma^2 - \sigma^2$$

$$= \sigma^2 - \frac{1}{n}\sigma^2 - \sigma^2 \quad = -\frac{1}{n}\sigma^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \bar{X})^2 = \frac{(n-1)}{n}\left[\frac{1}{(n-1)}\sum(X_i - \bar{X})^2\right] = \frac{(n-1)}{n}S^2$$

$Var(\hat{\sigma}^2) = Var\left[\frac{(n-1)}{n}S^2\right] = \frac{(n-1)^2}{n^2}\ V(S^2) = \frac{(n-1)^2}{n^2}\frac{2\sigma^4}{(n-1)}$

$= \frac{2(n-1)\sigma^4}{n^2}$

$$\lim_{n\to\infty} B(\hat{\sigma}^2) = \lim_{n\to\infty} \frac{-\sigma^2}{n} = 0$$

$$\lim_{n\to\infty} Var(\hat{\sigma}^2) = \lim_{n\to\infty} \frac{2(n-1)(\sigma^4)}{n^2} = 0$$

∴ MSE = Var $(\hat{\theta})$ + $\left[B\left(\hat{\theta}\right)\right]^2$

∴ $\lim_{n\to\infty} E\ (\hat{\sigma}^2 - \sigma^2)^2 = \lim_{n\to\infty} Var\ (\hat{\sigma})^2 + \lim_{n\to\infty}[B\ (\hat{\sigma}^2)]^2$

= 0 + 0 = 0

$\hat{\sigma}^2$ is consistent estimator of σ^2.

Sufficiency **الكفاية**

In the statistical inference problems on a parameter, one of the major questions is: Can a specific statistic replace the entire data without losing pertinent information?

في مشاكل الاستدلال الإحصائي على معلمة ، يكون أحد الأسئلة الرئيسية هو: هل يمكن لإحصاء محدد أن يحل محل البيانات بأكملها دون فقدان المعلومات ذات الصلة.

Suppose X_1 ,…, X_n is random sample from a probability distribution with unknown parameter θ. In general, statisticians look for ways of reducing a set of data so that these data can be more easily understood without losing the meaning associated with the entire collection of observations. Intuitively, a statistic U is a sufficient statistic for a parameter θ if U contains all the information available in the data about the value of θ.

For example, the sample mean may contain all the relevant information about the parameter μ, and in that case U = $\bar{X}$ is called a sufficient statistic for μ . An estimator that is a function of a sufficient statistic can be deemed to be a "good" estimator, because it depends on fewer data values. When we have a sufficient statistic U for , we need to concentrate only on U because it exhausts all the information that the sample has about θ. That is, knowledge of the actual n observations does not contribute anything more to the inference about θ.

Definition :

Let X_1,…., X_n be a random sample from a probability distribution with unknown parameter θ .Then, the statistic U=g (X_1,…., X_n) is said sufficient for θ. if the conditional pdf or pf of X_1,…., X_n given U = u does not depend on θ for any value of u. An estimator of U that is a function of a sufficient statistic for θ is said to be a sufficient estimator of θ.

Definition: **Simple consistency**

Let $T_1, T_2, \dots, T_n$ be a sequence of estimators of $\tau(\boldsymbol{\theta})$, where $T_n = t_n$ (X_1, ….,X_n). The sequence $\{T_n\}$ is defined to be a simple (or weakly) consistent sequence of estimators of τ ($\boldsymbol{\theta}$) if for every ε > 0 the following is satisfied:

$$\lim_{n\to\infty} P_{\boldsymbol{\theta}}[\tau(\theta) - \varepsilon < T_n < \tau(\boldsymbol{\theta}) + \varepsilon]$$

<u>Remark:</u> If an estimator is a mean-squared-error consistent estimator, it is also a simple consistent estimator, but not necessarily vice versa.

<u>Proof</u> :

$P_{\boldsymbol{\theta}}[\tau(\boldsymbol{\theta}) - \varepsilon < T_n < \tau(\boldsymbol{\theta}) + \varepsilon]$= P $[|T_n - \tau(\boldsymbol{\theta})| < \varepsilon]$

= $P_{\boldsymbol{\theta}}$ $[[T_n - \tau(\boldsymbol{\theta})]^2 < \varepsilon^2] \geq 1 - \frac{\mathcal{S}_{\boldsymbol{\theta}}[[T_n - \tau(\boldsymbol{\theta})]^2]}{\varepsilon^2}$

by the Chebyshev inequality. As n approaches infinity,$\mathcal{S}_{\boldsymbol{\theta}}$ $[[T_n - \tau(\boldsymbol{\theta})]^2]$ approaches 0. Hence lim $P_{\boldsymbol{\theta}}$ $[\tau(\boldsymbol{\theta}) - \varepsilon < T_n, < \tau(\boldsymbol{\theta}) + \varepsilon] = 1$

<u>Example:</u>

Let $x_1, \dots \dots, x_n$ be iid Bernoulli random variables with parameter θ. show that$\sum_{i=1}^{n} x_i$ is sufficient for θ.

<u>Solution:</u>

The joint probability mass function of $x_1, \dots \dots, x_n$ is

$$f(x_1, \dots \dots, x_n; \theta) = \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i}$$

Because $U = \sum_{i=1}^{n} x_i$ we have $f(x_1, \dots \dots, x_n; \theta) = \theta^U (1-\theta)^{n-U}$,$0 \leq U \leq n$.

Also, because $U \sim b(n, \theta)$ we have

$$f(u, \theta) = \binom{n}{u} \theta^U (1-\theta)^{n-U} \qquad ,0 \leq U \leq n$$

Also,

$$f(x_1, \dots, x_n | U = u) = \frac{f(x_1, \dots, x_n; u)}{f_U(u)} = \begin{cases} \frac{f(x_1, \dots, x_n)}{f_U(u)} & u = \sum_{i=1}^{n} x_i \\ 0 & o.w. \end{cases}$$

Therefore ,

$$f(x_1, \dots, x_n | U = u) = \frac{f(x_1, \dots, x_n; u)}{f_U(u)} = \begin{cases} \dfrac{\theta^u (1-\theta)^{n-u}}{\binom{n}{u} \theta^u (1-\theta)^{n-u}} = \dfrac{1}{\binom{n}{u}}, \text{for } \ u = \sum_{i=1}^{n} x_i \\ 0 \qquad\qquad\qquad\qquad o.w. \end{cases}$$

Which is independent of . Therefore U is sufficient for θ .

Example:

let $x_1, \dots \dots, x_n$ be arandom sample from passion (λ) show that the mean $\bar{x}$ is consistent to λ.

Solution:

$x_i \sim$ piosson Distribution

$$v(\bar{x}) = v\left[\sum \frac{x_i}{n}\right] \Rightarrow \frac{1}{n^2} v\left[\sum x_i\right] = \frac{1}{n^2} v[x_1 + x_2 + \dots + x_n]$$
$$= \frac{1}{n^2}[\lambda + \lambda + \cdots] == \frac{1}{n^2} n\lambda$$

$v(\bar{x}) = \frac{\lambda}{n}$ where $\epsilon = k\frac{\sigma}{n} = k\sqrt{\frac{\lambda}{n}} \quad \Rightarrow \quad k = \frac{\epsilon\sqrt{n}}{\sqrt{\lambda}} \quad \Rightarrow \quad k^2 = \frac{\epsilon^2\, n}{\lambda}$

$$P\left\{|\bar{x} - \lambda| > k\sqrt{\frac{\lambda}{n}}\right\} \le \frac{1}{\frac{\epsilon^2\, n}{\lambda}} = \frac{\lambda}{\epsilon^2\, n}$$

$$\lim_{n \to \infty} P\left\{|\bar{x} - \lambda| > k\sqrt{\frac{\lambda}{n}}\right\} \le \frac{\lambda}{\epsilon^2\, n} \quad by\ chebysheos = 0$$

$\lim_{n\to\infty}\left[\frac{\lambda}{\epsilon^2\, n}\right] = \frac{1}{\infty} = 0$ then $\bar{x}$ is consistent to λ

Example:

let $x_1, \dots \dots, x_n$ be a random sample from $N(\mu, \sigma)$, show that ${S_n}^2$ is consistent to σ^2, where ${S_n}^2 = \sum\left[\frac{x_i - \bar{x}}{n-1}\right]^2$.

Solution: Since $\frac{(n-1)}{\sigma^2}{S_n}^2 \sim \chi^2_{(n-1)}$ then $v(S^2) = 2r \quad since\ r = n-1$

$$v\left({S_n}^2\right) = 2(n-1)$$

$$v\left[\frac{n-1}{\sigma^2}{S_n}^2\right] = 2(n-1)$$

$$\left[\frac{(n-1)^2}{\sigma^4}v\left({S_n}^2\right) = 2\ (n-1)\right] * \frac{\sigma^4}{(n-1)^2}$$

$v(S^2) = \frac{2(n-1)\sigma^4}{(n-1)^2} \Rightarrow v(S^2) = \frac{2\sigma^4}{(n-1)}$ where $\epsilon = k\sigma_{S_n}$

$$\epsilon = k\sqrt{\frac{2\sigma^4}{(n-1)}} \quad \Rightarrow \quad k = \frac{\epsilon\sqrt{n-1}}{\sqrt{2\sigma^2}} \quad \Rightarrow \quad k^2 = \frac{\epsilon^2\ (n-1)}{2\sigma^4}$$

$$\lim_{n\to\infty}\left\{\left|{S_n}^2 - \sigma^2\right| > k\sqrt{\frac{2\sigma^4}{(n-1)}}\right\} \le \frac{1}{\frac{\epsilon^2\ (n-1)}{2\sigma^4}}$$

$$\lim_{n\to\infty}\left\{\left|{S_n}^2 - \sigma^2\right| > k\sqrt{\frac{2\sigma^4}{(n-1)}}\right\} \le \frac{2\sigma^4}{\epsilon^2\ (n-1)} \qquad by\ chebysheos = 0$$

$$\frac{2\sigma^4}{\infty} = 0$$

${S_n}^2$ is consistent to σ^2.

Republic of Iraq Ministry of Higher Education & Research

University of Anbar

College of Education for Pure Sciences

محاضرات الاحصاء ١

مدرس المادة : الاستاذ المساعد الدكتور فراس شاكر محمود

Roa – Black well theorem

The following theorem says that if we want an estimator with small MSE we can confine our search to estimators which are functions of the sufficient statistic.

Theorem 3.3 (Rao-Blackwell Theorem) *Let $\hat{\theta}$ be an estimator of θ with $\mathbb{E}(\hat{\theta}^2) < \infty$ for all θ. Suppose that T is sufficient for θ, and let $\theta^* = \mathbb{E}(\hat{\theta} \mid T)$. Then for all θ,*

$$\mathbb{E}(\theta^* - \theta)^2 \leq \mathbb{E}(\hat{\theta} - \theta)^2.$$

The inequality is strict unless $\hat{\theta}$ is a function of T.

Proof.

$$\mathbb{E}[\theta^* - \theta]^2$$

$$= \mathbb{E}\left[\mathbb{E}(\hat{\theta} \mid T) - \theta\right]^2 = \mathbb{E}\left[\mathbb{E}(\hat{\theta} - \theta \mid T)\right]^2 \leq \mathbb{E}\left[\mathbb{E}((\hat{\theta} - \theta)^2 \mid T)\right] = \mathbb{E}(\hat{\theta} - \theta)^2$$

The outer expectation is being taken with respect to T. The inequality follows from the fact that for any RV, W, $\text{var}(W) = \mathbb{E}W^2 - (\mathbb{E}W)^2 \geq 0$. We put $W = (\hat{\theta} - \theta \mid T)$ and note that there is equality only if $\text{var}(W) = 0$, i.e., $\hat{\theta} - \theta$ can take just one value for each value of T, or in other words, $\hat{\theta}$ is a function of T. ∎

If $\hat{\theta}$ is unbiased estimator for θ and t(x) is sufficient for θ, then the estimation $\bar{\theta}$ where

$$\bar{\theta} = E\left[\frac{\hat{\theta}}{t(x)}\right]$$

is also unbiased and its variance less than or equal to the variance of $\hat{\theta}$ i .e :

$$v(\bar{\theta}) \leq v(\hat{\theta})$$

Example : Let $x_1, x_2, \dots, x_n$ is ar.s from Ber(θ) if x_1 is unbiased est for θ, Find a better estimator by using the Roa_Black well Theorem .

<u>Solution :</u>

$$f(x,\theta) = \theta^x(1-\theta)^{1-x}$$

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f(x_i, \theta)$$

$$= \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}.1$$

h(x) does not depend upon θ. Then $\sum x_i$ is s.s for θ. We have $E(x_1) = \theta$. By using the Roa – Black well theorem, we get $\bar{\theta} = E\left(\frac{\hat{\theta}}{t(x)}\right) = E\left(\frac{x_1}{\sum x_i}\right)$ is better estimator than x_1. Now what is $\bar{\theta}$

$$E\left(\frac{x_1}{\sum x_i}\right) = E\left[\frac{x_1 = x_i}{\sum x_i = t}\right]$$

$x_1 = x_i$ is unbiased and $\sum x_i = t$ is sufficient

$$= \sum_{x=0,1} x_1 P\left[\frac{x_1 = x_i}{\sum x_i = t}\right]$$

$$= 0.P\left[\frac{x_1 = 0}{\sum x_i = t}\right] + 1.P\left[\frac{x_1 = 1}{\sum x_i = t}\right]$$

$$= P\left[\frac{x_1 = 1}{\sum x_i = t}\right] = \frac{P[x_1 ,\ \sum_{i=1}^{n} x_i = t]}{p(\sum_{i=1}^{n} x_i = t)}$$

$$P[x_1 = 1] = \theta$$

$$x \sim Ber(\theta)$$

$$\sum x_i \sim Binomal\ (n, \theta)$$

$$p\left(\sum x_i = t\right) = \binom{n}{t}\theta^t(1-\theta)^{n-t}$$

$$\sum x_i \sim Binomal(n-1, \theta)$$

$$P\left[\sum x_i = t-1\right] = \binom{n-1}{t-1}\theta^{t-1}(1-\theta)^{n-t}$$

$$\rightarrow \frac{P[x_1 = 1] * P[\sum_{i=2}^{n} x_i = t-1]}{P[\sum x_i = t]}$$

$$= \frac{\theta \binom{n-1}{t-1} \theta^{t-1}(1-\theta)^{n-t}}{\binom{n}{t} \theta^{t}(1-\theta)^{n-t}}$$

$$= \frac{\binom{n-1}{t-1}}{\binom{n}{t}} = \frac{\frac{(n-1)!}{(t-1)!\left((n-1)-(t-1)\right)!}}{\frac{n!}{t!\,(n-t)!}}$$

$$\frac{(n-1)!}{(t-1)!\,(n-t)!} - \frac{t!\,(n-t)!}{n!}$$

$$\frac{(n-1)!}{(t-1)!} * \frac{t_1(t-1)!}{n(n-1)!} = \frac{t}{n} = \frac{\sum x_i}{n} = \bar{x}$$

$\therefore\ \bar{x}$ is a better estimator than x_1 for θ.

Example:

Let $x_1, x_2, \dots, x_n \sim P(\theta) \ \rightarrow iid$ use the Roa – Black well theorem to find an estimator for θ better than x_1 iid= identically independent distribution

Solution:

$$f(x_1, \theta) = \frac{\theta^{x} e^{-\theta}}{x!}$$

$$f(x_1, x_2, \dots, x_n) = \frac{n\theta^{\sum x_i} e^{-n\theta}}{\prod_{i=1}^{n} x_i!} = n\theta^{\sum x_i} e^{-n\theta} * \frac{1}{\prod_{i=1}^{n} x_i!}$$

$\therefore \sum x_i$ is sufficient statistic (s.s.) for θ. Now , to Find $\bar{\theta}$

$$\bar{\theta} = E\left[\frac{x_1}{t(x)} = \sum_{i=1}^{n} x_i = t\right]$$

$$P\left[x_1 = \frac{X_1}{\sum x_i = t}\right] = \frac{P[x_1 = X_1, \sum_{i=1}^{n} x_i = t]}{P[\sum_{i=1}^{n} x_i = t]}$$

$$= \frac{p[x_1 = X_1, \sum x_i = t - x_1]}{p[\sum x_i = t]} = \frac{p[x_1 = X_1, \sum x_i = t]}{P[\sum x_i = t]}$$

$$P[x_1 = X_1] = \frac{\theta^{x_1} e^{-\theta}}{x_1!}$$

$$x \sim p(\theta)$$

$$\sum x_i \sim P(n\theta), \sum x_i \sim P((n-1)\theta)$$

$$P\left[\sum x_i = t\right] = \frac{(n\theta)^t e^{-n\theta}}{t!}$$

$$P\left[\sum x_i, t - x_1\right] = \frac{\left((n-1)\theta\right)^{t-x_1}}{(t-x_1)!}$$

$$P\left[\frac{x_1 = X_1}{\sum x_i = t}\right] = \frac{\frac{\theta^{x_1} e^{-\theta}}{x_i!} \; \frac{(n-1)\theta^{t-x_1} e^{-(n-1)\theta}}{(t-x)!}}{\frac{(n\theta)^t \, e^{-n\theta}}{t!}}$$

$$\frac{\theta^{x_1} e^{-\theta}}{x_1!} * \frac{(n-1)\theta^{t-x_1} e^{-(n-1)\theta}}{(t-x)!} * \frac{t!}{(n\theta)^t \, e^{-n\theta}}$$

$$\frac{t!\,(n-1)^{t-x_1}}{x_1!\,(t-x_1)!\,n^t}$$

Now $n^t = n^{x_1} * n^{t-x_1}$

$$P\left[\frac{x_1 = X_1}{\sum x_i = t}\right] = \frac{t_i(n-1)^{t-x_1}}{x_1!\,(t-x_1)!\,n^{x_1} * n^{t-x_1}}$$

$$= \frac{t!}{x_1!\,(t-x_1)!} * \left(\frac{1}{n}\right)^{x_1} \left(\frac{n-1}{n}\right)^{t-x_1}$$

$$= \frac{t!}{x_1!\,(t-x_1)!} \left(\frac{1}{n}\right)^{x_1} \left(1 - \frac{1}{n}\right)^{t-x_1}$$

$$\binom{t}{x_1} \left(\frac{1}{n}\right)^{x_1} \left(1 - \frac{1}{n}\right)^{t-x_1} \sim Bin\left(t, \frac{1}{n}\right)$$

$$E\left[\frac{x_1}{\sum x_i}\right] = t * \frac{1}{n} = \frac{t}{n} = \frac{\sum x_i}{n} = \bar{x}$$

$\bar{x}$ is a better estimator for θ

C**ompleteness :-** A statistic t(x) is said to be complete if for all θ the function h(t) statistic $E(h(t)) = 0$ which implies that $h(T) = 0$

Example: Let $x \sim Ber(\theta)$. Show that x is complete.

Solution :

$$f(x,\theta) = \theta^x (1-\theta)^{1-x}$$

We have

$$E(h(x) = 0 \text{ we prove } h(x) = 0$$

$$E\big(h(x)\big) = \sum_{x=0,1} h(x) * f(x,\theta) = 0$$

$$h(0) * f(0,\theta) + h(1) * f(1,\theta) = 0$$

$$h(0) * (1-\theta) + h(1) * \theta = 0$$

$$h(0) - h(0) * \theta + h(1) * \theta = 0$$

$$h(0) + \theta\big(h(1) - h(0)\big) = 0$$

$$\theta \neq 0 \text{ is perameter}$$

$$h(1) - h(0) = 0 \rightarrow h(1) = h(0)$$

$$\because h(0) = 0$$

$$h(1) = 0 \qquad x = 0,1$$

$$\therefore x\ is\ complete$$

Example:

Let $x_1, x_2, \dots, x_n$ is ar.s from a dist $Ber(\theta)$. Show that $T = \sum x_1$is complete sufficient statistic for θ

Solution:

$$f(x,\theta) = \theta^x (1-\theta)^{1-x}$$

$$\prod_{i=1}^{n} f(x,\theta) = \theta^{\sum x_i} (1-\theta)^{n-\sum x_i} * 1$$

$$\therefore \sum x_i\ is\ s.s\ for\ \theta$$

Now , we went to prove $T = \sum x_i$ is C.S.S

$$x \sim Ber(\theta)$$

$$\sum x_i \sim Binonal\ (n, \theta)$$

$$E\big(h(t)\big) = 0$$

$$E[h(T)] = \sum_{T=o}^{n} h(T) * f(T, \theta) = 0$$

$$f(T, \theta) = f\left(T = \sum x_i, \theta\right) = \binom{n}{T} * \theta^T (1-\theta)^{n-T}$$

$$E\big(h(T)\big) = h(0)\binom{n}{0}\theta^0(1-\theta)^{n-0} + h(1)\binom{n}{1}\theta(1-\theta)^{n-1} + \cdots$$

$$+h(n)\binom{n}{n}\theta^n(1-\theta)^{n-n} = 0$$

$$h(0)(1-\theta)^n = 0\ \ \%(1-\theta)^n$$

$$h(0) = 0$$

$$h(1)\binom{n}{1}\theta(1-\theta)^{n-1} = 0\ \ \%\binom{n}{1}\theta(1-\theta)^{n-1}$$

$$h(1) = 0$$

$$h(0) = h(1) = \cdots = h(n) = 0$$

$$h(T) = 0\ , T = 1,2, \dots, n$$

$$\sum x_i\ \ is\ C.S.S\ for\ \theta$$

Exponential Family of distribution

<u>Definition</u>: A one Parameter exponential family of distribution is that if $f(x, \theta)$ can be express in the from

$$f(x, \theta) = a(\theta) * b(x) e^{c(\theta)dx}$$

$$or\ f(x, \theta) = e^{c(\theta)dx} + b(x) + a(\theta) \quad \alpha < x < \beta$$

Where α , β does not depot upon θ.

Example: if $x \sim Ber(\theta)$, Show that $f(x,\theta)$. belongs to exponential family

$$f(x,\theta) = \theta^x(1-\theta)^{1-x}$$

$$= \theta^x(1-\theta)(1-\theta)^{-x}$$

$$= \theta^x(1-\theta) * \frac{1}{(1-\theta)^x}$$

$$= (1-\theta)\left(\frac{\theta}{1-\theta}\right)^x$$

$$= (1-\theta)e^{\ln\left(\frac{\theta}{1-\theta}\right)^x}$$

$$= (1-\theta)e^{\mathrm{x}\ln\left(\frac{\theta}{1-\theta}\right)}$$

$$a(\theta) = (1-\theta), b(x) = 1, c(\theta) = \ln\left(\frac{\theta}{1-\theta}\right), d(x) = x$$

$f(x,\theta)$ belongs to exponential family.

H.w : $x \sim P(\theta)$ show that $f(x,\theta)$ belong to exponential family .

Theorem :

Let $f(x,\theta)$ be a P.d.f which represent a regular case of the exponential class . Than if $x_1, x_2, \dots, x_n$. Where (n) is a fixed positive integer is a random sample from a distribution , with P. d .f $f(x,\theta)$ the statistic $t = \sum_{i=1}^{n} di$ is sufficient statistic for θ and the family $g(t,\theta)$ of probability density family of t is complete that is t is C.S.S for θ.

Theorem : Any function of C.S.S is MVUE of it expectation

Example: if $x \sim \exp(\theta)$ find MVUE

$$f(x,\theta) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$$

$$a(\theta) = \frac{1}{\theta} \ , b(x) = 1 \ , c(\theta) = -\frac{1}{\theta} \ , d(x) = x$$

$\mathrm{f(x,\theta)} =$ belong to exponential family $\mathrm{t} = \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{d(x_i)} = \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{x_i}$ is C.S.S for θ .