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

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

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

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

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

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

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,

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