LECTURE NOTE

ON

PROBABILITY AND SATISTICS 2

BY

ASSIST. PRF. DR. MUSTAFA I. NAIF

DEPARTMENT OF MATHEMATICS
COLLEGE OF EDUCATION FOR PURE SCIENCE
UNIVERISTY OF ANBAR

LECTURE 5#

➢ Outline :-

✓ Discrete distributions

4- Uniform distribution

Definition

Expected value and Variance

Moment generating function

Characteristic function

Distribution function

Solved exercises

Exercises

Uniform distribution

Definition : *A random variable* X *has a* ***discrete uniform distribution and it is referred to as a discrete uniform random variable if and only if its probability mass function is given by:***

$$f(x) = \frac{1}{k} \text{ for } x = 1, 2, \ldots, k,$$

We denoted by: (X~DU(k))

Proof: You do that.

Expectation and Variance:

$$\sum_{x=1}^{k} x = \frac{k(k+1)}{2}$$

$$E(X) = \sum_{x=1}^{k} x\, f(x) = \sum_{x=1}^{k} x \frac{1}{k} = \frac{1}{k} \sum_{x=1}^{k} x$$ **, So,** $$\boxed{E(X) = \frac{(k+1)}{2}}$$

Def. of Var.

$$\boxed{Var(X) = \frac{k^2 - 1}{12}}$$ Tp proof that you have to use

$$\sum_{x=1}^{k} x^2 = \frac{k(k+1)(2k+1)}{6}$$

Uniform distribution

Moment generating and Characteristic function :

If X is a r.v. distributed as a discrete uniform dist., then the m.g.f. of X is given as follows:

Geometric series

$$M_X = E(e^{tX}) = \frac{1}{k}\sum_{x=1}^{k} e^{tX} = \frac{1}{k}\sum_{x=1}^{k} Z^x, \qquad Z = e^t$$

$$= \frac{1}{k}(Z + Z^2 + \cdots + Z^k) = \frac{Z}{k}(1 + Z + Z^2 + \cdots + Z^{k-1})$$

But $\sum_{x=0}^{k-1} Z^x = \frac{1-Z^k}{1-Z}$, then $M_X = \frac{Z}{k} . \frac{1-Z^k}{1-Z} = \frac{e^t(1-e^{kt})}{k(1-e^t)} = \boxed{\frac{e^t(e^{kt}-1)}{k(e^t-1)}}$; t>0

By the same way, we can get the characteristic function as follows:

$$\varphi_X(t) = \boxed{\frac{e^{it}(e^{kit}-1)}{k(e^{it}-1)}};$$

Uniform distribution

Distribution function: The distribution function of a discrete uniform random variable X is:

$$F(X) = P(X \leq x) = \sum_{u=1}^{x} f(u) = \sum_{u=1}^{x} \frac{1}{k} = \boxed{\frac{x}{k}} \quad ; \quad x = 1,2,\ldots,k$$

Example1: Let X~DU(8). Find pmf, CDF, E(X), Var(X) and P(X ≤4).

Sol.: f(x)= $\frac{1}{8}$, F(x) $= \frac{x}{8}, E(X) = 4.5, Var(X) = \frac{63}{12}$

$P(X \leq 4) = F(4) = 0.5$. (Try to find $P(X \geq 3)$?).

Example2: Let X~DU(k).Find the mean and the variance of Y=a+bX where a and be are two real constants.

Sol.: It will be direct by using the properties of discrete uniform distribution.

5-Hypergeometric Distribution

Consider a collection of n objects which can be classified into two classes, say class 1 and class 2. Suppose that there are n_1 objects in class 1 and n_2 objects in class 2. A collection of r objects is selected from these n objects at random and without replacement. We are interested in finding out the probability that exactly x of these r objects are from class 1. If x of these r objects are from class 1, then the remaining r − x objects must be from class 2. We can select x objects from class 1 in any one of $\binom{n_1}{x}$ways. Similarly, the remaining r − x objects can be selected in $\binom{n_2}{r-x}$ ways. Thus, the number of ways one can select a subset of r objects from a set of n objects, such that x number of objects will be from class 1 and r − x number of objects will be from class 2, is given by $\binom{n_1}{x}$ $\binom{n_2}{r-x}$ Hence,

$$P(X = x) = \frac{\binom{n_1}{x}\binom{n_2}{r-x}}{\binom{n}{r}},$$

where $x \leq r$, $x \leq n_1$ and $r - x \leq n_2$.

Hypergeometric Distribution

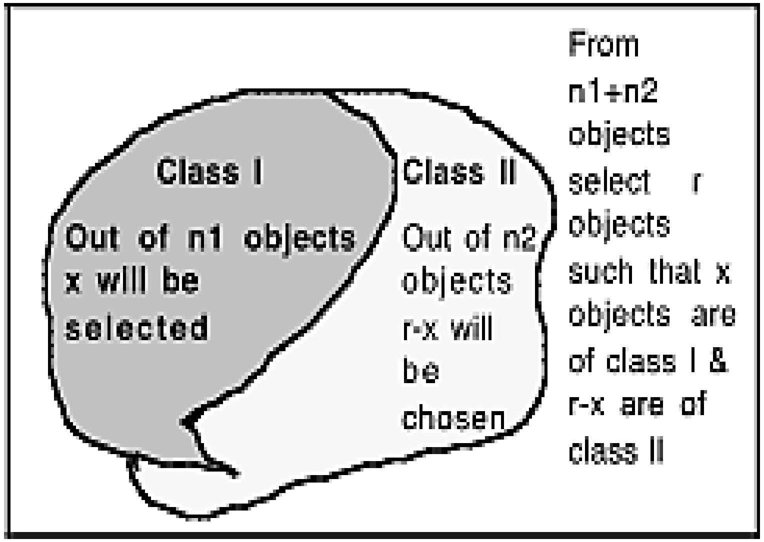
Definition : A random variable X is said to have a hypergeometric distribution if its probability mass function is of the form:

$$f(x) = \frac{\binom{n_1}{x}\binom{n_2}{r-x}}{\binom{n_1+n_2}{r}}, \qquad x = 0, 1, 2, \ldots, r$$

where $x \leq n_1$ and $r - x \leq n_2$ with n_1 and n_2 being two positive integers.

We shall denote such a random variable by writing $X \sim HYP(n_1, n_2, r)$.

Example :Suppose there are 3 defective items in a lot of 50 items. A sample of size 10 is taken at random and without replacement. Let X denote the number of defective items in the sample. What is the probability that the sample contains at most one defective item?



Hypergeometric Distribution

Answer: Clearly, $X \sim HYP(3, 47, 10)$. Hence the probability that the sample contains at most one defective item is

$$\begin{aligned} P(X \leq 1) &= P(X=0) + P(X=1) \\ &= \frac{\binom{3}{0}\binom{47}{10}}{\binom{50}{10}} + \frac{\binom{3}{1}\binom{47}{9}}{\binom{50}{10}} \\ &= 0.504 + 0.4 \\ &= 0.904. \end{aligned}$$

Theorem If $X \sim HYP(n_1, n_2, r)$, then

$$E(X) = r\,\frac{n_1}{n_1+n_2}$$

$$Var(X) = r\left(\frac{n_1}{n_1+n_2}\right)\left(\frac{n_2}{n_1+n_2}\right)\left(\frac{n_1+n_2-r}{n_1+n_2-1}\right).$$

Hypergeometric Distribution

Proof: Let $X \sim HYP(n_1, n_2, r)$. We compute the mean and variance of X by computing the first and the second factorial moments of the random variable X. First, we compute the first factorial moment (which is same as the expected value) of X. The expected value of X is given by

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{r} x\, f(x) \\
&= \sum_{x=0}^{r} x \frac{\binom{n_1}{x}\binom{n_2}{r-x}}{\binom{n_1+n_2}{r}} \\
&= n_1 \sum_{x=1}^{r} \frac{(n_1-1)!}{(x-1)!\,(n_1-x)!} \frac{\binom{n_2}{r-x}}{\binom{n_1+n_2}{r}} \\
&= n_1 \sum_{x=1}^{r} \frac{\binom{n_1-1}{x-1}\binom{n_2}{r-x}}{\frac{n_1+n_2}{r}\binom{n_1+n_2-1}{r-1}} \\
&= r\,\frac{n_1}{n_1+n_2} \sum_{y=0}^{r-1} \frac{\binom{n_1-1}{y}\binom{n_2}{r-1-y}}{\binom{n_1+n_2-1}{r-1}}, \qquad \text{where } y = x-1 \\
&= r\,\frac{n_1}{n_1+n_2}.
\end{aligned}
$$

The last equality is obtained since $\sum_{y=0}^{r-1} \frac{\binom{n_1-1}{y}\binom{n_2}{r-1-y}}{\binom{n_1+n_2-1}{r-1}} = 1.$ where $\sum_{i=0}^{n}\binom{a}{i}\binom{b}{n-i} = \binom{a+b}{n}$

Similarly, we find the second factorial moment of X to be

$E(X(X-1)) = \dfrac{r(r-1)\,n_1\,(n_1-1)}{(n_1+n_2)\,(n_1+n_2-1)}$. Therefore, the variance of X is

$$\begin{aligned} Var(X) &= E(X^2) - E(X)^2 \\ &= E(X(X-1)) + E(X) - E(X)^2 \\ &= \frac{r(r-1)\,n_1\,(n_1-1)}{(n_1+n_2)\,(n_1+n_2-1)} + r\,\frac{n_1}{n_1+n_2} - \left(r\,\frac{n_1}{n_1+n_2}\right)^2 \\ &= r\left(\frac{n_1}{n_1+n_2}\right)\left(\frac{n_2}{n_1+n_2}\right)\left(\frac{n_1+n_2-r}{n_1+n_2-1}\right). \end{aligned}$$

Distribution Function: The distribution function of a discrete hypergeometric random variable X is:

$F(X) = P(X \le x) = \sum_{k=c}^{x} \frac{\binom{n_1}{x}\binom{n_2}{r-x}}{\binom{n_1+n_2}{r}}$, where c=max(0,r-$n_1 + n_2$)

Moment generating function :

The m g. f. of a discrete hypergeometric random variable X is:

$$M_X(t) = \frac{(n_1 - r)!\,(n_1 - n_2)!}{n_1}.H(-r; -n_2; n_1 - n_2 + 1; e^t)$$

where $H(-r; -n_2; n_1 - n_2 + 1; e^t)$=$\sum_{j=0}^{\infty} \frac{(-r)^{[j]}(-n_2)^{[j]}(e^t)^j}{(n_1-n_2-r+1)^{[j]}j!}$ and in general ,

for any number a , then :

$$a^{[j]} = a(a+1)(a+2)\ldots(a+j-1).$$

Note: Let X1, X2 are r.v's distributed as Ber(p). If X2 is not independent of X1, and we should not expect X to have a binomial distribution. (why?)

See you next Lecture