Module 28

SOME SPECIAL DISCRETE DISTRIBUTIONS

- X : a discrete r.v. with support S_X, d.f. $F_X(\cdot)$ and p.m.f. $f_X(\cdot)$
- $\mu = E(X) = \sum_{x \in S_X} x f_X(x)$
- $\sigma^2 = \mathrm{Var}(X) = E((X-\mu)^2) = \sum_{x \in S_X} (x-\mu)^2 f_X(x)$
- For any function $h(\cdot)$

$$E(h(X)) = \sum_{x \in S_X} h(x) f_X(x),$$

  provided the sum is finite.
- Moment generating function

$$\mathrm{M}_X(t) = E(e^{tX}) = \sum_{x \in S_X} e^{tx} f_X(x)$$

I. Bernoulli and Binomial Distribution

Bernoulli Experiment: A random experiment that results in just two possible outcomes, say success (S) and failure (F).

- Then
$$\Omega = \{S, F\},\ \mathcal{P}(\Omega) = \{\phi, \Omega, \{S\}, \{F\}\}.$$
- Let $P(\{S\}) = p$ and $P(\{F\}) = 1 - p = q$ (say), where $0 < p < 1$.
- Define $X : \Omega \to \mathbb{R}$
$$X(\omega) = \begin{cases} 1, & \text{if } \omega = S \\ 0, & \text{if } \omega = F \end{cases}.$$
$$= \text{ No. of successes in single trial.}$$
- The support of X is $S_X = \{0, 1\}$ and p.m.f. of X is
$$f_X(x) = P(\{X = x\}) = \begin{cases} q, & \text{if } x = 0 \\ p & \text{if } x = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} p^x q^{1-x}, & \text{if } x = 0, 1 \\ 0, & \text{otherwise} \end{cases}.$$

$\rightarrow$ Bernoulli distribution with success probability $p \in (0, 1)$.

- Now consider a sequence of n (where n is a fixed positive integer) independent Bernoulli trials, with probability of success in each trial as $p \in (0, 1)$.
- Define

$$X = \text{ No. of successes in } n \text{ Bernoulli trials.}$$

- Then $S_X = \{0, 1, 2, \ldots, n\}$ and p.m.f. of X is

$$\begin{aligned} f_X(x) &= P(\{X = x\}) \\ &= P(\{\text{S in } x \text{ trials and F in } n - x \text{ trials}\}) \\ &= \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & \text{if } x = 0, 1, \ldots, n \\ 0, & \text{otherwise} \end{cases}. \end{aligned}$$

$\rightarrow$ Binomial distribution with parameters n and p (written as $X \sim \mathrm{Bin}(n, p)$; $n \in \mathbb{N}$ and $p \in (0, 1)$ are parameters).

- A $\mathrm{Bin}(1, p)$ distribution is a Bernoulli distribution with success probability $p \in (0, 1)$.
- Suppose that $X \sim \mathrm{Bin}(n, p)$, where $n \in \mathbb{N}$ and $p \in (0, 1)$. Let $q = 1 - p$. Then

$$\begin{aligned}
\mathrm{M}_X(t) &= E(e^{tX}) \\
&= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x q^{n-x} \\
&= (q + pe^t)^n, \;\; t \in \mathbb{R}
\end{aligned}$$

- $$\begin{aligned} \mathrm{M}_X^{(1)}(t) &= npe^t(q+pe^t)^{n-1},\ t\in\mathbb{R} \\ \mathrm{M}_X^{(2)}(t) &= npe^t(q+pe^t)^{n-1}+n(n-1)p^2e^{2t}(q+pe^t)^{n-2},\ t\in\mathbb{R} \end{aligned}$$

- $$\begin{aligned} \mu = E(X) &= \mathrm{M}_X^{(1)}(0) = np \\ E(X^2) &= \mathrm{M}_X^{(2)}(0) = np + n(n-1)p^2 \\ \sigma^2 = \mathrm{Var}(X) &= E(X^2) - (E(X))^2 \\ &= np(1-p) \\ &= npq \end{aligned}$$

- Mean $>$ Variance

Result 1 :

Let $X_1, \ldots, X_k$ be independent r.v.s with $X_i \sim \mathrm{Bin}(n_i, p)$; $n_i \in \mathbb{N}$, $p \in (0,1)$, $i = 1, \ldots, k$. Then $Y = \sum_{i=1}^{k} X_i \sim \mathrm{Bin}(\sum_{i=1}^{k} n_i, p)$.

Proof. For $t \in \mathbb{R}$

$$\begin{aligned}
\mathrm{M}_Y(t) &= E\left(e^{t\sum_{i=1}^{k} X_i}\right) \\
&= \prod_{i=1}^{k} \mathrm{M}_{X_i}(t) \\
&= \prod_{i=1}^{k} (q + pe^t)^{n_i} \\
&= (q + pe^t)^{\sum_{i=1}^{k} n_i}
\end{aligned}$$

Now the result follows by uniqueness of m.g.f.

II. Negative Binomial distribution

- Consider a sequence of independent Bernoulli trials with probability of success in each trial as $p \in (0,1)$. Let $r \in \mathbb{N}$ be a fixed positive integer.
- Define $Y =$ No. of failures preceding the rth success, so that $Y + 1$ is the number of trials required to get the rth success.
- Then $S_Y = \{0, 1, 2, \ldots\}$ and, for $y \in S_Y$,

$$
\begin{aligned}
f_Y(y) &= P(\{Y = y\}) \\
&= P(y \text{ failures precede } r\text{th success}) \\
&= P(r-1 \text{ successes in first } y+r-1 \text{ trial, and } (y+r)\text{th} \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{trial is success}) \\
&= \binom{y+r-1}{r-1} p^{r-1}(1-p)^y p \\
&= \binom{y+r-1}{r-1} p^r (1-p)^y
\end{aligned}
$$

- Thus the p.m.f. of Y is

$$f_Y(y) = \begin{cases} \binom{y+r-1}{r-1} p^r q^y, & \text{if } y = 0, 1, 2, \ldots \\ 0, & \text{otherwise} \end{cases}.$$

$\rightarrow$ Negative binomial distribution with parameters $r \in \mathbb{N}$ and $p \in (0,1)$ (written as $X \sim \mathrm{NB}(r,p)$).

- Note that

$$\begin{aligned} \sum_{y \in S_Y} f_Y(y) &= p^r \sum_{y=0}^{\infty} \binom{y+r-1}{r-1} q^y \\ &= p^r \left[1 + rq + \frac{r(r+1)}{2!} q^2 + \frac{r(r+1)(r+2)}{3!} q^3 + \ldots \right] \\ &= p^r (1-q)^{-r} \\ &= 1. \end{aligned}$$

- The m.g.f. of Y is

$$\begin{aligned} \mathrm{M}_Y(t) &= E(e^{tY}) \\ &= \sum_{y=0}^{\infty} e^{ty} \binom{y+r-1}{r-1} p^r q^y \\ &= p^r \sum_{y=0}^{\infty} \binom{y+r-1}{r-1} (qe^t)^y \\ &= p^r (1-qe^t)^{-r}, \quad t < -\ln q \end{aligned}$$

-

$$\begin{aligned} \psi_Y(t) &= \ln \mathrm{M}_Y(t) \\ &= r \ln p - r \ln(1-qe^t) \end{aligned}$$

-

$$\begin{aligned}
\psi_Y^{(1)}(t) &= \frac{rqe^t}{1-qe^t}, \quad t < -\ln q \\
\psi_Y^{(2)}(t) &= rq\frac{(1-qe^t)e^t + qe^{2t}}{(1-qe^t)^2}, \quad t < -\ln q \\
\mu = E(Y) = \psi_Y^{(1)}(0) &= \frac{rq}{p} \\
\sigma^2 = \mathrm{Var}(Y) = \psi_Y^{(2)}(0) &= \frac{rq}{p^2};
\end{aligned}$$

- Mean $<$ Variance

Result 2 :

Let $Y_1, Y_2 \ldots, Y_k$ be independent r.v.s with $Y_i \sim \mathrm{NB}(r_i, p)$; $i = 1, \ldots, k$.

Then $Y = \sum\limits_{i=1}^{k} Y_i \sim \mathrm{NB}\big(\sum\limits_{i=1}^{k} r_i, p\big)$.

Proof. For $t < -\ln q$

$$\begin{aligned} \mathrm{M}_Y(t) &= \prod_{i=1}^{k} \mathrm{M}_{Y_i}(t) \\ &= \prod_{i=1}^{k} \left(\frac{p}{1-qe^t}\right)^{r_i} \\ &= \left(\frac{p}{1-qe^t}\right)^{\sum\limits_{i=1}^{k} r_i} \end{aligned}$$

Now the result follows by using uniqueness of m.g.f.

- An $\mathrm{NB}(1, p)$ distribution is called a geometric distribution (distribution of number of failures preceding first success) and is denoted by $\mathrm{Ge}(p),\ 0 < p < 1$. If $Y \sim \mathrm{Ge}(p),\ 0 < p < 1$, then

$$P(Y \geq m) = \sum_{y=m}^{\infty} p(1-p)^y = (1-p)^m,\ m = 1, 2, \ldots$$

$$P(Y \geq j + k | Y \geq j) = (1-p)^k = P(Y \geq k),\ j, k \in \{0, 1, 2, \ldots\}$$

$\rightarrow$ Lack of memory property (Interpret it when Y represents the lifetime of an item)

III. The Hypergeometric distribution

- Consider a population having N objects out of which a are marked and $N-a$ are unmarked. A random sample of size n is drawn from this population without replacement.
- Let $X =$ No. of marked objects in the sample of n objects.
- Then

$$\begin{aligned} S_X &= \{x \in \mathbb{N} : 0 \leq x \leq n, 0 \leq x \leq a, n-x \leq N-a\} \\ &= \{x \in \mathbb{N} : \max\{0, n-N+a\} \leq x \leq \min\{n, a\}\} \end{aligned}$$

$$\begin{aligned} f_X(x) &= P(\{X = x\}) \\ &= \begin{cases} \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}, & \text{if } x \in S_X \\ 0, & \text{otherwise} \end{cases}. \end{aligned}$$

$\rightarrow$ Hypergeometic distribution ($X \sim \mathrm{Hyp}(a, n, N)$, $N \in \{1, 2, \ldots\}, a \in \{1, 2, \ldots, N\}, n \in \{1, 2, \ldots, N\}$).

- Since the support S_X is finite, it follows that the m.g.f. $\mathrm{M}_X(t)$ is finite for every $t \in \mathbb{R}$, although a closed form expression for it can not be obtained.
- Let $\psi_r(X) = X(X-1)\ldots(X-r+1)$, $r = 1, 2, \ldots$. Then it can be shown that, for $r = 1, 2, \ldots$,

$$E(\psi_r(X)) = \begin{cases} \frac{\binom{N-r}{n-r}}{\binom{N}{n}} a(a-1)\ldots(a-r+1), & \text{if } r \leq \min\{n, a\} \\ 0, & \text{otherwise} \end{cases}.$$

- In particular

$$E(X) = n\frac{a}{N}$$

and, for $n \geq 2,\ a \geq 2$

$$E(X(X-1)) = n(n-1)\frac{a(a-1)}{N(N-1)}$$

$$\begin{aligned}
\mathrm{Var}(X) &= E(X^2) - (E(X))^2 \\
&= E(X(X-1)) + E(X) - (E(X))^2 \\
&= n\left(\frac{a}{N}\right)\left(1 - \frac{a}{N}\right)\frac{N-n}{N-1}.
\end{aligned}$$

IV. The Poisson Distribution

A r.v. X is said to have the Poisson distribution with parameter $\lambda > 0$ (written as $X \sim \mathrm{P}(\lambda)$) if its p.m.f. is given by

$$f_X(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & \text{if } x = 0, 1, 2, \ldots \\ 0, & \text{otherwise} \end{cases}.$$

- Clearly $S_X = \{0, 1, 2, \ldots\}$, $f_X(x) \geq 0, \forall\; x \in \mathbb{R}$ and $\sum_{x \in S_X} f_X(x) = 1$.
- For $t \in \mathbb{R}$,

$$\begin{aligned} \mathrm{M}_X(t) &= E(e^{tX}) \\ &= \sum_{x \in S_X} e^{tx} f_X(x) \\ &= \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} \end{aligned}$$

$$
\begin{aligned}
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\
&= e^{-\lambda} e^{\lambda e^t} \\
&= e^{\lambda(e^t - 1)}.
\end{aligned}
$$

-

$$
\begin{aligned}
\psi_X(t) &= \ln \mathrm{M}_X(t) = \lambda(e^t - 1), \;\; t \in \mathbb{R} \\
\psi_X^{(1)}(t) &= \lambda e^t, \;\; t \in \mathbb{R} \\
\psi_X^{(2)}(t) &= \lambda e^t, \;\; t \in \mathbb{R}
\end{aligned}
$$

-

$$
\mu = E(X) = \psi_X^{(1)}(0) = \lambda, \qquad \sigma^2 = \mathrm{Var}(X) = \psi_X^{(2)}(0) = \lambda
$$

- Mean = Variance.

Result 3 :

Let $X_1, X_2, \ldots, X_k$ be independent random variables with $X_i \sim \mathrm{P}(\lambda_i),\ \lambda_i > 0, i = 1, \ldots, k$. Then $Y = \sum\limits_{i=1}^{k} X_i \sim \mathrm{P}(\sum\limits_{i=1}^{k} \lambda_i)$.

Proof. For $t \in \mathbb{R}$

$$\begin{aligned} \mathrm{M}_Y(t) &= E(e^{tY}) \\ &= \prod_{i=1}^{k} \mathrm{M}_{X_i}(t) \\ &= \prod_{i=1}^{k} e^{\lambda_i(e^t-1)} \\ &= e^{\left(\sum\limits_{i=1}^{k} \lambda_i\right)(e^t-1)} \end{aligned}$$

which is the m.g.f. of $\mathrm{P}\left(\sum\limits_{i=1}^{k} \lambda_i\right)$ distribution. Now the result follows by uniqueness of m.g.f.s.

V. The Discrete Uniform Distribution

Let $N \geq 1$ be given integer. A r.v. X is said to follow uniform distribution on $\{1, 2, \ldots, N\}$ (written as $X \sim \mathrm{U}(1 - N)$) if its p.m.f. is given by

$$\begin{aligned} f_X(x) &= P(X = x) \\ &= \begin{cases} \frac{1}{N}, & \text{if } x \in \{1, 2, \ldots, N\} \\ 0, & \text{otherwise} \end{cases}. \end{aligned}$$

- Clearly $S_X = \{0, 1, 2, \ldots, N\}$,

$$\mathrm{M}_X(t) = \frac{1}{N} \sum_{x=1}^{N} e^{tx}$$

$$\mu = E(X) = \sum_{x \in S_X} x f_X(x) = \frac{1}{N} \sum_{x=1}^{N} x = \frac{N+1}{2}$$

$$E(X^2) = \sum_{x \in S_X} x^2 f_X(x) = \frac{1}{N} \sum_{x=1}^{N} x^2 = \frac{(N+1)(2N+1)}{6}$$

$$\begin{aligned} \sigma^2 = \mathrm{Var}(X) &= E(X^2) - (E(X))^2 \\ &= \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} \\ &= \frac{N^2-1}{12} \end{aligned}$$

Take Home Problems

(1) (a) Suppose that, for some $n \in \mathbb{N}$ and $p \in (0,1)$, $X \sim \mathrm{Bin}(n,p)$. Show that $Y = n - X \sim \mathrm{Bin}(n, 1-p)$.
    (b) If $X \sim \mathrm{Bin}(n, \frac{1}{2})$, show that the distribution of X is symmetric. Hence find $P(X \leq \frac{n}{2})$.

(2) Let X be a discrete type r.v. with support $S_X = \{0, 1, 2, \ldots\}$. Show that the probability distribution of X has lack of memory property if and only if $X \sim \mathrm{Ge}(p)$, for some $p \in (0,1)$.

(3) Suppose that $X \sim \mathrm{Hyp}(a, n, N)$, where $a, n, N \in \mathbb{N}, a \leq N$ and $n \leq N$. Find the value of

$$E(X(X-1)\ldots(X-r+1)),$$

where $r \in \mathbb{N}$.

Thank you for your patience