DISCRETE DISTRIBUTIONS

A r.v X is said to have a discrete Distribution if the range of X, the sample space, is countable.

① Discrete uniform Distribution:

$$X \sim unif(1, N)$$

$$\mathbb{P}(X = x \mid N) = \frac{1}{N} \qquad x = 1, 2, \dots, N$$

identities:

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2} \quad \text{and} \quad \sum_{i=1}^{k} i^2 = \frac{k(k+1)(2k+1)}{6}$$

$$\mathbb{E}[X] = \sum_{x=1}^{N} x \cdot \mathbb{P}(X = x \mid N)$$

$$= \sum_{x=1}^{N} x \cdot \frac{1}{N} = \frac{1}{\cancel{N}} \frac{\cancel{N}(N+1)}{2}$$

$$\mathbb{E}[X] = \frac{N+1}{2}$$

$$\mathbb{E}[x^2] = \sum_{x=1}^{N} x^2 \, \mathbb{P}(X = x \mid N)$$

$$= \frac{1}{N} \sum_{x=1}^{N} x^2$$

$$= \frac{1}{\cancel{N}} \cdot \frac{\cancel{N}(N+1)(2N+1)}{6}$$

$$\mathbb{E}[x^2] = \frac{(N+1)(2N+1)}{6}$$

$$\mathrm{Var}(x) = \mathbb{E}[x^2] - \mathbb{E}[x]^2$$

$$= \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4}$$

$$= \frac{(N+1)(4N+2 - 3N-3)}{12}$$

$$\mathrm{Var}(x) = \frac{(N+1)(N-1)}{12}$$

$$\Rightarrow \boxed{\begin{aligned} \mathbb{E}[x] &= \frac{N+1}{2} \\ \mathrm{Var}(x) &= \frac{(N+1)(N-1)}{12} \end{aligned}}$$

2 Hypergeometric Distribution:

- many application in finite population sampling.
- Example : urn model.

* we have Large urn of N Ball's that are identical in every way, except M are red and $N-M$ are green.

N Ball's = identical M red Ball's + identical $N-M$ green Ball's

* we reach in, Blindfolded, and select K ball's at random.

(K ball's are taken at once, unordered)

* $\mathbb{P}$(exactly x of the Ball's are red) ?

$$\mathbb{P}(X = x \mid N, M, K)$$

$$= \frac{\binom{M}{x} \cdot \binom{N-M}{K-x}}{\binom{N}{K}} \qquad x = 0, 1, 2, \ldots K$$

$$\sum_{x=0}^{K} \mathbb{P}(X = x) = \sum_{x=0}^{K} \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}$$

$$= 1$$

$$\mathbb{E}[X] = \sum_{x=0}^{K} x \cdot \mathbb{P}(X = x)$$

$$= \sum_{x=1}^{K} x \cdot \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}$$

for $x = 0$

the value $= 0$

$$= \sum_{x=1}^{K} \frac{M \binom{M-1}{x-1} \binom{N-M}{K-x}}{\frac{N}{K}\binom{N-1}{K-1}}$$

$$= \frac{KM}{N} \sum_{x=1}^{K} \frac{\binom{M-1}{x-1}\binom{N-M}{K-x}}{\binom{N-1}{K-1}}$$

$$y = x-1$$

$$= \frac{KM}{N} \underbrace{\sum_{y=0}^{K-1} \frac{\binom{M-1}{y}\binom{N-M}{(K-1)-y}}{\binom{N-1}{K-1}}}_{\text{Pdf} = 1}$$

$$\boxed{\mathbb{E}[X] = \frac{KM}{N}}$$

$$E[X(X-1)] = \sum_{x=2}^{K} x(x-1) \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}$$

$$= \sum_{x=2}^{K} \frac{M(M-1)\binom{M-2}{x-2}\binom{N-M}{K-x}}{\frac{N\,N-1}{K\,(K-1)}\binom{N-2}{K-2}}$$

$$\mathbb{E}[X(X-1)] = \frac{M(M-1)K(K-1)}{N(N-1)}$$

$$\operatorname{var}(X) = \mathbb{E}[X(X-1)] + E[X] - E[X]^2$$

$$= \frac{M(M-1)K(K-1)}{N(N-1)} + \frac{KM}{N} - \frac{K^2M^2}{N^2}$$

$$= KM\left[\frac{N(M-1)(K-1) + N(N-1) - KM(N-1)}{N^2(N-1)}\right]$$

$$= \frac{KM}{N}\left(\frac{\cancel{NMK} - NM - NK + \cancel{N} + N^2 - \cancel{N} - \cancel{NMK} + KM}{N(N-1)}\right)$$

$$= \frac{KM}{N}\left(\frac{KM + N^2 - NM - NK}{N(N-1)}\right)$$

$$\boxed{\operatorname{var}(X) = \frac{KM}{N}\left(\frac{(N-M)(N-K)}{N(N-1)}\right)}$$

3 Binomial distribution

$$X \sim \text{Bernoulli}(p)$$

$$X = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$$

$$\Rightarrow \quad \mathbb{E}[X] = p$$

$$Var(X) = p(1-p)$$

$$\mathbb{P}(X = x) = p^{x} (1-p)^{1-x} \quad x \in \{0,1\}$$

* n identical Bernoulli trails

Y = # number of success

$$\mathbb{P}(Y = y \mid n, p) = \binom{n}{y} p^{y} (1-p)^{n-y}$$

$$y = 0, 1, 2 \ldots n$$

Theorem Binomial theorem

$$(x+y)^{n} = \sum_{i=0}^{n} \binom{n}{i} x^{i} y^{n-i}$$

$$\mathbb{E}[X] = np$$

$$Var(X) = np(1-p)$$

(4) Poission Distribution:

— Serve as a model for a number of different types of experiments. (waiting for an occurrence), (number of occurance in a given amount of time)

$$\mathbb{P}(X = x \mid \lambda) = \frac{e^{-\lambda} \lambda^{x}}{x!} \qquad x = 0, 1, \ldots$$

λ = intensity parameter.

we know that the taylor expansion of e^y

$$e^y = \sum_{i=0}^{\infty} \frac{y^i}{i!}$$

$$\Rightarrow \sum_{x=0}^{\infty} \mathbb{P}(X=x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^{i}}{i!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{i}}{i!}$$

$$= e^{-\lambda} \cdot e^{\lambda} = 1$$

$$\mathbb{E}[X] = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda} \lambda^x}{x!}$$

$$= \sum_{x=1}^{\infty} \frac{e^{-\lambda} \cdot \lambda^{x-1}}{(x-1)!} \cdot \lambda$$

$$= \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} = \lambda$$

$$\mathbb{E}[X(X-1)] = \sum_{x=0}^{\infty} \frac{x(x-1)\, e^{-\lambda} \lambda^x}{x!}$$

$$= \sum_{x=2}^{\infty} \frac{e^{-\lambda} \lambda^{x-2}}{(x-2)!} \cdot \lambda^2$$

$$= \lambda^2$$

$$\Rightarrow \quad \text{var}(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2$$

$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

$$\boxed{\begin{aligned} \mathbb{E}[X] &= \lambda \\ \text{var}(X) &= \lambda \end{aligned}}$$

$$M_X(t) = \mathbb{E}[e^{tX}]$$

$$= \sum_{x=0}^{\infty} \frac{e^{tx} e^{-\lambda} \lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} \frac{e^{-\lambda} (\lambda e^{t})^{x}}{x!}$$

$$= \frac{e^{-\lambda}}{e^{-\lambda e^t}} \sum_{x=0}^{\infty} \frac{e^{-\lambda e^t} (\lambda e^t)^{x}}{x!}$$

$$\boxed{M_X(t) = e^{\lambda(e^t - 1)}}$$

$$\mathbb{P}(X=x) = \frac{\lambda}{x} \cdot \mathbb{P}(X=x-1)$$

Ex: Consider a telephone operator who, on the average, handle 5 calls every 3 min.

$\mathbb{P}$(No call's in the next min)?

$\mathbb{P}$ (at least 2 calls)?

Solution: $\lambda = \mathbb{E}[X] = \frac{5}{3}$

$$\mathbb{P}(X=0) = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-5/3}$$

$$\mathbb{P}(X \geq 2) = 1 - P(X=0) - P(X=1)$$

$$= 1 - e^{-5/3} - \frac{5}{3}e^{-5/3}$$

$$= 1 - e^{-5/3} - \frac{5}{3}e^{-5/3}$$

Example 3.2.5 (Poisson approximation)

A typesetter, on the average, makes one error in every 500 words. A typical page contain's 300 words. What is the Prob that there will be no more than two error's in five pages?

Binomial $p = \frac{1}{500}$ $n = 1500$

$$\mathbb{P}(X \leq 2) = \sum_{x=0}^{2} \binom{1500}{x} \left(\frac{1}{500}\right)^x \left(\frac{499}{500}\right)^{1500-x}$$

$= 0.4230$

<u>Poisson</u> :-

one error in every 500 words

$\Rightarrow \quad \lambda = 1$ per 500 words

$\Rightarrow \quad \lambda = 3$ per 1500 words

$$\Rightarrow \mathbb{P}(x \le 2) = \mathbb{P}(x=0) + \mathbb{P}(x=1) + \mathbb{P}(x=2)$$

$$= e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda}{2} \cdot \lambda \cdot e^{-\lambda}$$

$$= e^{-3}\left(1 + 3 + \frac{9}{2}\right)$$

$$= e^{-3}\left(\frac{17}{2}\right) = 0.42319$$

5

Negative Binomial Distribution

(Pascal)

Number of Bernoulli trails required to get a fixed number of successes.

X = # of trails required for r success

= $(r-1)$ successes in $x-1$ trails,

r^{th} success in x^{th} trail.

$$\mathbb{P}(X=x \mid r,p) = \binom{x-1}{r-1} p^{r-1} \cdot (1-p)^{x-r} \cdot p$$

$x = r, r+1, \ldots$

$X \sim$ negative binomial (r,p)

Y = number of failures before the r^{th} success

$\Rightarrow$ X = number of trailes before r^{th} success

$$Y = X - r$$

$$\mathbb{P}(X = x) = \binom{x-1}{r-1} p^{r-1} (1-p)^{x-r} \cdot p$$

$$x = r, r+1, r+2 \ldots$$

$$\mathbb{P}(Y = y) = \binom{y+r-1}{r-1} p^{r-1} (1-p)^{y} p$$

$$\mathbb{P}(Y = y) = \binom{y+r-1}{r-1} p^{r} (1-p)^{y}$$

$$y = 0, 1, \ldots$$

$$\binom{r+y-1}{r} = \frac{(r+y-1)}{(r-1)!\, y!}$$

$$= \frac{(r+y-1)(r+y-2)\ldots\ldots(r+2)(r+1)\, r \cdot \cancel{(r-1)!}}{\cancel{(r-1)!}\, y!}$$

$$= (-1)^{y} \frac{(-r)(-r-1)\ldots(-r-y+1)}{y!}$$

$$\mathbb{P}(Y=y) = (-1)^y \binom{-r}{y} p^r (1-p)^y$$

$$\mathbb{E}[Y] = \sum_{y=0}^{\infty} y \binom{r+y-1}{y} p^r (1-p)^y$$

$$= \sum_{y=1}^{\infty} \frac{y\,(r+y-1)!}{(r-1)!\,y!} p^r (1-p)^y$$

$$= \sum_{y=1}^{\infty} \frac{(r+y-1)!}{r!\,(y-1)!} \cdot p^{r+1} \cdot (1-p)^{y-1} \left(\frac{1-p}{p}\right) \cdot r$$

$$= \left(\frac{1-p}{p}\right) r \underbrace{\sum_{y=1}^{\infty} \binom{r+y-1}{y-1} \cdot p^{r+1} \cdot (1-p)^{y-1}}_{\text{pmf with } (r+1,\, y-1) \;=\; 1}$$

$$\mathbb{E}\{Y\} = \left(\frac{1-p}{p}\right) r$$

$$\mathbb{E}[y(y-1)]$$

$$= \sum_{y=0}^{\infty} y(y-1) \frac{(r+y-1)!}{(r-1)!\, y!} p^{r} (1-p)^{y}$$

$$= \sum_{y=2}^{\infty} \frac{(r+y-1)!}{(r+1)!\,(y-2)!} p^{r+2} (1-p)^{y-2} \cdot \frac{(1-p)^2}{p^2} \cdot r(r+1)$$

$$\mathbb{E}[Y(Y-1)] = \frac{(1-p)^2}{p^2} \cdot r(r+1)$$

$$\operatorname{Var}(Y) = \mathbb{E}[Y(Y-1)] + E[Y] - E[Y]^2$$

$$= \frac{(1-p)^2}{p^2} \cdot r(r+1) + r\frac{(1-p)}{p} - r^2\left(\frac{1-p}{p}\right)^2$$

$$= \frac{(1-p)r\left((1-p)(r+1) + p - r(1-p)\right]}{p^2}$$

$$= \frac{(1-p)r\left(1 - p\cancel{r} + \cancel{r} - \cancel{p} + \cancel{p} - \cancel{r} + \cancel{p}r\right]}{p^2}$$

$$= \frac{(1-p)r}{p^2}$$

$$\boxed{\mathbb{E}[Y] = r\frac{(1-p)}{p}, \quad \operatorname{Var}(Y) = \frac{r(1-p)}{p^2}}$$

Example 3.2.6 (Inverse binomial sampling)

if the proportion of individual possessing a certain characteristic is p and we sample until we see r such individual's, the the number of individuals sampled is a negative binomial random variable

Ex: $r = 100$

$$\mathbb{P}(X \geq N) = \sum_{x=N}^{\infty} \binom{x-1}{99} p^{100} (1-p)^{x-100}$$

$$= 1 - \sum_{x=100}^{N-1} \binom{x-1}{99} p^{100} \cdot (1-p)^{x-100}$$

6

Geometric Distribution:

negative binomial distribution with $r=1$

$\Rightarrow$ waiting time for the rth occurance

$$\mathbb{P}(X = x \mid p) = p(1-p)^{x-1} \qquad x = 1, 2, \ldots$$

X can be interpretted as the trail at which the first success occurs, so we are "waiting for a success".

$$\sum_{x=1}^{\infty} \mathbb{P}(X=x \mid p) = \sum_{x=1}^{\infty} p(1-p)^{x-1}$$

$$= p \sum_{x=1}^{\infty} (1-p)^{x-1}$$

$$= p \sum_{y=0}^{\infty} (1-p)^{y}$$

$$= p \cdot \frac{1}{1-(1-p)} = 1$$

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} p(1-p)^{x-1} \cdot x$$

$$= p \sum_{x=1}^{\infty} (1-p)^{x-1} \cdot x = \frac{p}{p^2} = \frac{1}{p}$$

or

$X = Y + 1$ where Y = negative binomi

$$\mathbb{E}[X] = \mathbb{E}[Y] + 1$$

$$= \frac{1-p}{p} + 1 = \frac{1}{p}$$

$$\mathrm{var}(X) = \mathrm{var}(Y+1) = \mathrm{var}(Y)$$

$$= \frac{1-p}{p^2}$$

Geometric distribution has a interesting property called Memoryless.

$$\forall s > t, \quad \mathbb{P}(X > s \mid X > t) = P(X > s-t)$$

$\Rightarrow$ geometric distribution "forget's" what occured.