Discrete Random Variables (II)

Suppose X is a discrete random variable

$\mathcal{X}$ = set of values of X. → finite / countably infinite.

Today's Goal:

Study some discrete r.v's which have $\mathcal{X}$ to be infinite.

① Poisson Distribution

X is said to have Poisson Distribution with parameter: $\lambda > 0$ → if $\mathcal{X} = \{0,1,2,3,\dots\}$

for $x \in \mathcal{X}$

$$\boxed{p_X(x) = \frac{e^{-\lambda}\cdot\lambda^x}{x!}}$$

Note: ①

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots = \sum_{x=0}^{\infty}\frac{\lambda^x}{x!}$$

$$\Rightarrow \boxed{1 = e^{-\lambda}\cdot\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = \sum_{x=0}^{\infty} e^{-\lambda}\cdot\frac{\lambda^x}{x!} = \sum_{x=0}^{\infty} p_X(x)}$$

Also, $\boxed{p_X(x) = e^{-\lambda}\cdot\frac{\lambda^x}{x!} \geq 0 \quad \forall x}$

∴ p_X is a legitimate probability Mass Function.

② If X has the Poisson Distribution with parameter λ → $E(X) = \mu_X = \lambda$

$V(X) = \sigma_X^2 = \lambda$

∴ $\sigma_X = \sqrt{\lambda}$

③ The Poisson distribution is often used to model phenomena where we are waiting for events to happen.

⇕

For example: In a "Poisson Process" → the number of occurrences in a given time interval t.

② Negative Binomial Distribution

X has the Negative Binomial Distribution with parameters: r, p → if $\mathcal{X} = \{0,1,2,3,\dots\}$

for $x \in \mathcal{X}$

$$\boxed{p_X(x) = \binom{x+r-1}{r-1} p^r (1-p)^x}$$

Note:

① The experiment for this distribution → Perform Bernoulli trials with $P(\text{success}) = p$ repeatedly until we get exactly r-successes

X = # of failures that precede the r^{th} success.

Then

$$P(X=x) = P(\text{exactly } x \text{ failures before the } r^{th} \text{ success})$$

Total ways we can have exactly x failures before r^{th} success → Choosing x spots in $x+r-1$ possible places (since the last spot has to be a success)

⇓

can be done in $\binom{x+r-1}{x} = \binom{x+r-1}{r-1}$ ways.

$$\therefore P(X=x) = \binom{x+r-1}{r-1} \underbrace{p^r}_{\text{since } r \text{ successes}} \underbrace{(1-p)^x}_{\text{since } x \text{ failures}}$$

② If X has negative Binomial Distribution with parameters: r, p → $E(X) = \mu_X = \frac{r(1-p)}{p}$

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

③ Called negative binomial

$$\boxed{\binom{x+r-1}{x} = (-1)^x\binom{-r}{x} = (-1)^x\frac{(-r)(-r-1)(-r-2)\cdots(-r-x+1)}{x\cdot(x-1)(x-2)\cdots 3\cdot 2\cdot 1}}$$

So that

$$\boxed{p_X(x) = (-1)^x\binom{-r}{x}p^r(1-p)^x}$$

⇓

very similar to the pmf of Binomial Distribution.

③ Geometric Distribution

X has Geometric Distribution with parameter: p → if $\mathcal{X} = \{1,2,3,\dots\}$

for $x \in \mathcal{X}$

$$\boxed{p_X(x) = p(1-p)^{x-1}}$$

Note:

① Geometric distribution → special case of Negative Binomial pmf

⇓

when $r=1$ and $x = 1,2,3,\dots$

② Is the simplest of the "waiting time" distributions

⇓

X can be interpreted as the trial at which first success occurs

⇓

"waiting for a success"

③ If X has Geometric distribution with "success" parameter: p → $E(X) = \mu_X = \frac{1}{p}$

$V(X) = \sigma_X^2 = \frac{(1-p)}{p^2}$

④ Sometimes used to model "lifetime" or "time until failure" of components.

Relationships Between Discrete distributions (in the limit)

① Binomial ↔ Hypergeometric

Let $N, M, n \in \mathbb{N}$, $p = \frac{M}{N}$

(i) $bin(x; n, p)$ → pmf for binomial with params: n = sample size, p = prob of success

(ii) $hyper(x; N, M, n)$ → pmf for hypergeometric distribution with params: N = population size, M = # of successes, n = sample size

If $N, M \to \infty$ st $\frac{M}{N} \to p$ → $\{hyper(x; N, M, n) \to bin(x; n, p)$

⇕

if sample size n is small compared to population size N, can assume that samples are "approximately" independent.

② Binomial ↔ Poisson

Suppose

$bin(x; n, p)$ → pmf of Binomial with params: n = sample size, p = probability of success.

$pois(x; \lambda)$ → pmf of Poisson Distribution with parameter $\lambda > 0$.

If $n \to \infty$ and $p \to 0$ such that $n\cdot p \to \lambda$ → $\{bin(x; n, p) \to pois(x; \lambda)\}$

⇕

If n is large, and p very small

⇕

$bin(x; n, p) \approx pois(x, \lambda = n\cdot p)$

"The Poisson Distribution is approximately Binomial for rare events."

$b(x; 100, 0.001)$ $x = 20$; $P(X \geq x)$

$$bin(20; 100, 0.001) = \binom{100}{20}\cdot(0.001)^{20}\cdot(1-0.001)^{80}$$

$$\frac{100\times 99\times 98\times\cdots 81}{20!} \quad (\quad)$$

$$pois(x, \lambda) = e^{-\lambda}\cdot\frac{\lambda^x}{x!} \; ; \; pois(20; \lambda = np) = pois(20; 0.1) = e^{-0.1}\cdot\frac{(0.1)^{20}}{20!}$$

$X \sim pois(\lambda) \rightsquigarrow \mathcal{X} = \{0,1,2,\dots,\}$

$\lambda > 0$

$$p_X(x) = P(X=x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}$$

$$\begin{cases} p_X(x) \geq 0 \quad \forall x \in \mathcal{X} \\ \Updownarrow \\ e^{-\lambda} \cdot \frac{\lambda^x}{x!} \geq 0 \end{cases}$$

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots$$

$$= \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \qquad \text{(multiply both sides by } e^{-\lambda})$$

$$1 = e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \Rightarrow 1 = \sum_{x=0}^{\infty} \left(e^{-\lambda} \frac{\lambda^x}{x!} \right)$$

$\longrightarrow p_X(x)$ for Poisson distribution.

x	0	1	2	3	4	...
$P_X(x)$	$e^{-\lambda}$	$e^{-\lambda} \lambda$	$e^{-\lambda} \frac{\lambda^2}{2!}$	$e^{-\lambda} \frac{\lambda^3}{3!}$		

$$p_X(x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!} \quad ; \quad p_X(0) = e^{-\lambda}$$

$$p_X(1) = e^{-\lambda} \cdot \lambda$$

$$p_X(2) = e^{-\lambda} \cdot \frac{\lambda^2}{2!}$$

$$\vdots$$

$$p_X(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

$$E(X) = 0 \cdot p_X(0) + 1 \cdot p_X(1) + 2 \cdot p_X(2) + \dots$$

$$= \lambda$$

$$Var(X) = E(X^2) - (E(X))^2 = \lambda$$

Typically used to calculate the chance/prob of getting a certain # events in a given.

(Check out textbook for Poisson process with rate α)

Example:

Telephone operator on avg 5 calls over a 3 minute period.

What is the probability that there will be no calls in the next minute?

Assuming that the telephone calls follow a Poisson Process

$$P_k(t) = P\left(\text{getting } k \text{ calls over a time period } t\right)$$

$$\Updownarrow$$

Poisson distribution with rate $\lambda = \alpha \cdot t$

α: # of occurrences in unit time.

5 calls in 3 minutes

$\Rightarrow$ $5/3$ calls in a minute $\rightsquigarrow \alpha = \frac{5}{3}$

X = # of calls in a minute

$\Updownarrow$

Poisson dist with $\lambda = \frac{5}{3}$

$$P(X=0) = P(\text{no calls in a minute})$$

$$= e^{-5/3}$$

$$P(\text{at least 2 calls (inclusive) in one min period}) = P(X \geq 2)$$

$$= \sum_{x=2}^{\infty} P(X=x) = \sum_{x=2}^{\infty} p_X(x)$$

$$1 = \sum_{x=0}^{\infty} p_X(x) = p_X(0) + p_X(1) + \sum_{x=2}^{\infty} p_X(x)$$

$$p_X(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \qquad \lambda = 5/3$$

$$P(X \geq 2) = \sum_{x=2}^{\infty} p_X(x) = 1 - p_X(0) - p_X(1)$$

$$= 1 - e^{-5/3} - e^{-5/3} \cdot (5/3)$$

$$= 1 - e^{-5/3}\left(1 + \frac{5}{3}\right)$$

Y = # calls in 2 minutes.

$\alpha = \frac{5}{3} \qquad t = 2$

$\therefore \lambda = 2 \cdot \alpha = \frac{10}{3}$

$P\left(\text{0 calls in 2 minutes}\right) = P(Y=0)$

$= e^{-10/3}$

$e^{-5/3} > e^{-10/3}$

Negative Binomial Distribution

X has Neg Bin dist

r = # successes

p = prob of getting a success

$\rightsquigarrow$ $X = \{0, 1, 2, \dots\}$

$$p_X(x) = \binom{x+r-1}{x} \cdot p^r \cdot (1-p)^x$$

Inverse Binomial Sampling

Population $\rightsquigarrow$ sampling from this population

$\{$successes, failures$\}$ $\rightsquigarrow$ we are aware of $P(S) = p$.

X = # failures until the r^{th} success.

$P(X = N) = P\left(\text{of exactly } N \text{ failures before the } r^{th} \text{ success}\right)$

$P(X \geq N) = P\left(\text{at least } N \text{ failures before the } r^{th} \text{ success}\right)$

$$r, p \rightsquigarrow \sum_{x=N}^{\infty} p_X(x)$$

Geometric Distribution

X = # trials until a success

$\{1, 2, 3, 4, \dots\}$

$$p_X(x) = p \cdot (1-p)^{(x-1)}$$

p: first success.

$(1-p)^{(x-1)}$: $(x-1)$ failures

$$\text{nbin}(x; r, p) \longrightarrow \text{pois}(x; \lambda)$$

$$\Uparrow$$

$$\left\{\begin{array}{ll} r \to \infty & r \cdot (1-p) \to \lambda \\ p \to 1 & (1-p) \to 0 \end{array}\right\}$$

If r is large $p \to 1$; $(1-p) \to 0$.

$P\left(\text{getting } k \text{ failures until the } r^{th} \text{ success}\right) \sim \frac{e^{-\lambda} \cdot \lambda^k}{k!}$

$$= e^{-r(1-p)} \cdot \frac{(r(1-p))^k}{k!}$$

$$\binom{x+r-1}{x} \cdot p^r \cdot (1-p)^x$$