11/02/2017

Lecture16

$Var[X] = E[(X-\mu)^2]$

mean
pivot
expectation

$E[X^2] = \sigma^2 + \mu^2$

$\sigma^2 = E[X^2] - \mu^2$

$Var[aX+c] = a^2\sigma^2$

$SE[aX+c] = |a|\sigma$

$T_2 = X_1 + X_2$

$E[T_2] = E[X_1] + E[X_2]$

$Var[T_2] = ?$

$(X_1+X_2-\mu_1-\mu_2)(X_1+X_2-\mu_1-\mu_2)$

$Var[X_1+X_2] = E[((X_1+X_2) - (\mu_1+\mu_2))^2]$

$= E[X_1^2 + X_2^2 + \mu_1^2 + \mu_2^2 + 2X_1X_2 - 2X_1\mu_1 - 2X_1\mu_2 - 2X_2\mu_1 - 2X_2\mu_2 + 2\mu_1\mu_2]$

$= \underbrace{E[X_1^2]}_{\sigma_1^2+\mu_1^2} + \underbrace{E[X_2^2]}_{\sigma_2^2+\mu_2^2} + \cancel{\mu_1^2} + \cancel{\mu_2^2} + 2E[X_1X_2] - \cancel{2\mu_1^2} - 2\mu_1\mu_2 - \cancel{2\mu_2\mu_1} - \cancel{2\mu_2^2} + \cancel{2\mu_1\mu_2}$

$= \sigma_1^2 + \sigma_2^2 + 2(E[X_1X_2] - \mu_1\mu_2)$

only when independent

covariance
$Cov[X_1, X_2]$

$g(X_1, X_2)$

$E[X_1X_2] = \sum_{X_1}\sum_{X_2} x_1x_2\, p(X_1, X_2) = \sum_{X_1}\sum_{X_2} x_1x_2\, p(X_1)p(X_2) = \underbrace{\sum_{X_1} X_1 p(X_1)}_{\mu_1} \underbrace{\sum_{X_2} X_2 p(X_2)}_{\mu_2}$

Assume X_1, X_2 are independent $\Rightarrow p(X_1, X_2) = p(X_1)p(X_2)$

Covariance is 0 when independent

$Cov[X_1, X_2] = E[X_1 X_2] - \mu_1\mu_2 = \mu_1\mu_2 - \mu_1\mu_2 = 0$

$\Rightarrow \boxed{Var(X_1 + X_2) = \sigma_1^2 + \sigma_2^2 \text{ if } X_1 X_2 \text{ independent}}$

IF $X_1, \dots X_n$ independent

$Var[X_1 + \dots X_n] = \sum_{i=1}^{n} Var[X_i] \overset{\text{if iid}}{=} n\sigma^2$

$Var[\bar{X}_n] = Var[\frac{1}{n}T_n] = \frac{1}{n^2}Var[T_n] = \frac{1}{n^2}\sum Var[X_i] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n} \xrightarrow[n\to\infty]{} 0$

$E[\bar{X}_n] = E[\frac{1}{n}T_n] = \frac{1}{n}E[T_n] = \frac{1}{n}n\mu = \mu$

$\Rightarrow SE(\bar{X}_n) = \frac{\sigma}{\sqrt{n}} \xrightarrow[n\to\infty]{} 0$

$X \sim Bin(n,p)$ $\quad X = X_1 + \dots + X_n$ where $X_1 \dots X_n$ iid $Bern(p)$

$E[X] = np$ $\quad E[X] = \sum E[X_i] = np$

$Var[X] = E[(X-\mu)^2] = \sum_{x=0}^{n} (x-np)^2 \binom{n}{x} p^x (1-p)^{n-x}$

$Var[X] = \sum Var[X_i] = n\sigma^2 = np(1-p) = SE[X] = \boxed{\sqrt{np(1-p)}}$

$X \sim Geom(p) = (1-p)^{x-1}p$

$E[X] = \frac{1}{p} = \mu$

$Var[X] = E[X^2] - \mu^2$

$= E[X^2] - \frac{1}{p^2}$

$E[X^2] = \sum_{x=1}^{\infty} x^2 (1-p)^{x-1} p$

have to start at 1, need to get a success, if you do 0 trials won't be able to have any successful trials let $d = x-1 \Rightarrow x = d+1$

$$E[X^2] = \sum_{x=1}^{\infty} x^2 (1-p)^{x-1} p = \sum_{d=0}^{\infty} (d+1)^2 (1-p)^d p = \sum_{d=0}^{\infty} d^2 (1-p)^d p + 2\sum_{d=0}^{\infty} d(1-p)^d p + \sum_{d=0}^{\infty} (1-p)^d p$$

$$\underbrace{(1-p) \sum_{d=1}^{\infty} d^2 (1-p)^{d-1} p}_{\Sigma[X^2]} + \underbrace{2(1-p) \sum_{d=1}^{\infty} d(1-p)^{d-1} p}_{E[X] = \frac{1}{p}} + \underbrace{p \sum_{d=0}^{\infty} (1-p)^d}_{\underbrace{\frac{1}{p}}_{1}}$$

$$E[X^2] = (1-p)E(X^2) + \frac{2(1-p)}{p} + 1\frac{p}{p}$$

$$E[X^2] - (1-p)E[X^2] = \frac{2(1-p)+p}{p}$$

$$E[X^2]p = \frac{2(1-p)+p}{p}$$

$$E[X^2] = \frac{2(1-p)+p}{p^2} = \frac{2-2p+p}{p^2} = \frac{2-p}{p^2}$$

$$Var[X] = E[X^2] - \frac{1}{p^2}$$

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

$X \sim Hyper(n, K, N)$

$$Var[X] = \sum_{x \in supp(X)} \left(x - n\frac{K}{N}\right)^2 \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

Nightmare... not in this class

$X_1, X_2, \ldots \overset{iid}{\sim} \text{Bern}(p)$

$X \sim \text{Geom}(p)$ "stopping time"

$P(X=17) = (1-p)^{16} p$

$P(X=7) = (1-p)^{6} p$

$$P(X=7) \overset{?}{=} P(X=17 \mid X>10) = \frac{P(X=17 \text{ and } X>10)}{P(X>10)} = \frac{P(X=17)}{P(X>10)}$$

$P(X>10) = 1 - F(10)$

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

1 … 10 … 17

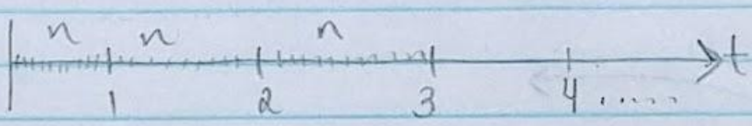
$$\frac{(1-p)^{16}p}{(1-p)^{10}} = P(X=7)$$

$$P(X=a+b \mid X>a) = \frac{P(X=a+b \text{ and } X>a)}{P(X>a)} = \frac{P(X=a+b)}{P(X>a)} = \frac{(1-p)^{a+b-1}p}{(1-p)^{a}} =$$

$$(1-p)^{b-1}p = P(X=b)$$

"memoryless property"

"it's like the first 10 did not happen"



$p(t) = (1-p)^{t-1}p$ within each time period, we run n $\text{Bern}(p)$ iid experiments

$\to F(t) = 1-(1-p)^t \to$ CDF

$p(t) = (1-p)^{nt-1}p$ "squeezing more and more experiments in"

let $\lambda = np \Rightarrow p = \frac{\lambda}{n}$ $\quad = \left(1-\frac{\lambda}{n}\right)^{nt-1}\frac{\lambda}{n}$ $\quad F(t) = 1-\left(1-\frac{\lambda}{n}\right)^{nt}$

END OF MIDTERM 2 MATERIAL // FINAL

PMF disappears when $n \to \underline{\underline{\infty}}$!

$\lim_{n\to\infty} p(t) =$

$\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{nt} \frac{\lambda}{n} = \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{nt} \lim_{n\to\infty} \frac{\lambda}{n} = 0 \;\forall_t$

limiting PMF

$\sum_{t=0}^{\infty} p(t) = 0$! "Broken"

not a PMF ! → Problem NOT A DISCRETE R.V

$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e^x$

$\lim_{n\to\infty} F(t) = \lim_{n\to\infty} 1-\left(1-\frac{\lambda}{n}\right)^{nt} = 1-\Bigg(\underbrace{\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n}}_{e^{-\lambda}}\Bigg)^t$

limiting CDF

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

$= 1 - \frac{1}{e^{\lambda t}}$ $t \to \infty = 1$

$t \to -\infty = 0$

legitimate CDF

$(\frac{2}{3})$