POPULATION VARIANCE : $\sigma^2 = \dfrac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$

SAMPLE VARIANCE : $s^2 = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$

or : $\dfrac{n\sum_{i=1}^{n}x_i^2 - (\sum_{i=1}^{n}x_i)}{n(n-1)}$

for any 2 sets A and B: $P(A\cup B) = P(A)+P(B)-P(A\cap B)$

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

$$P(A|B) = \frac{P(AB)}{P(B)}$$

Theorem of Total Probability : for an event A in the sample space S and $B_1, B_2, .., B_n$ be n mutually exclusive and exhaustive events in S

$$P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$$

Bayes theorem: for an event A in the sample space S and $B_1, B_2, .., B_n$ be n mutually exclusive and exhaustive events in S

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}$$

$$E(X) = \sum_{\text{all } i} x_i P(X=x_i) \quad \text{or} \quad \int_{-\infty}^{\infty} x f_X(x)\,dx$$

$$V(X) = E(X^2) - [E(X)]^2$$

$X \sim bin(n,p) \quad \Rightarrow P(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \quad ; k=0,1,\ldots,n$

$E(X) = np \quad V(X) = np(1-p)$

$X \sim geo(p) \quad \Rightarrow P(X=k) = q^{k-1} p \quad ; k=1,2,\ldots$

$E(X) = \frac{1}{p}$

$X \sim$ negative binomial $(r,p) \Rightarrow P(Y=k) = \left[\binom{k-1}{r-1} p^{r-1} (1-p)^{k-r}\right.$

$E(X) = \frac{r}{p}$ for $k = r, r+1, \ldots$

$X \sim$ Poisson$(\lambda) \Rightarrow P(X=k$ in t time periods$) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$

$E(X) = \lambda t$

for $k = 0,1,2,\ldots$ and $\lambda > 0$

$T \sim exp(\lambda) \quad \Rightarrow f_T(t) = \lambda e^{-\lambda t} \quad t \geq 0, \lambda > 0$

$E(T) = \frac{1}{\lambda} \quad V(T) = \frac{1}{\lambda^2}$

JOINT PROBABILITY DISTRIBUTIONS $P(X=x_i, Y=y_j)$ and $f_{X,Y}(x,y)$:

Marginal distributions: $P(X=x_i) = \sum_{\text{all } j} P(X=x_i, Y=y_j)$ for each i

$P(Y=y_j) = \sum_{\text{all } i} P(X=x_i, Y=y_j)$ for each j

also: $P_Y(y_j) = \sum_{\text{all } i} P_X(x_i) P_{Y|X}(y_j|x_i)$ for each j THM OF TOTAL PROBABILITY (another form)

Marginals: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$

EXPECTED VALUE: $E(X) = \sum_{\text{all } i} x_i P(X=x_i) \quad E(g(X)) = \sum_{\text{all } i} g(x_i) P(X=x_i)$

$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx \quad E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$

FORMULA SHEET for EXAM 4

- if $X \sim N(\mu, \sigma_x)$, then $\bar{X}_n \sim N\left(\mu, \frac{\sigma_x}{\sqrt{n}}\right)$

- if $X \sim \text{bin}(n, p)$, then $E(X) = np$ $\quad V(X) = np(1-p)$ $\quad \sigma_x = \sqrt{np(1-p)}$

  and if n large, X can be approximated by $N\left(\mu, \sqrt{np(1-p)}\right)$

- $X \sim \text{bin}(n, p)$ then for large n $\hat{p}$ (sample proportion) has the probability distribution:

$$\hat{p} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

- $(1-\alpha)\%$ C.I. for μ: $\quad \bar{X}_n \pm Z_{\frac{\alpha}{2}}\left(\frac{\sigma_x}{\sqrt{n}}\right)$ — can use s_x for σ_x if $n \geq 30$

- $(1-\alpha)\%$ C.I. for p: $\quad \hat{p} \pm Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$

- choosing sample size:
  - when estimating μ: $\quad E$ = maximum tolerable error $= Z_{\alpha/2}\frac{\sigma_x}{\sqrt{n}}$
  - when estimating p: $\quad E$ = " " " $= Z_{\alpha/2}\left(\sqrt{\frac{p(1-p)}{n}}\right)$

- small sample estimation for μ: $\quad \bar{X}_n \pm t_{n-1, \alpha/2}\left(\frac{s_x}{\sqrt{n}}\right)$

TO ESTIMATE (TEST) POPULATION MEAN: