STATS 1

WEEKS 1-4

① Variance (sample variance)

$$\frac{\sum_{i=0}^{n} (x_i - \mu)^2}{n-1} \quad ; \quad \mu \rightarrow \text{mean}$$

② Outlier

Upper bound $\rightarrow$ $Q_3 + 1.5(IQR)$
Lower bound $\rightarrow$ $Q_1 - 1.5(IQR)$
$IQR = Q_3 - Q_1$

③ Percentile

p = percentile ; n = no' of observations

$$n \cdot p = \begin{cases} \dfrac{(n \cdot p) + (n \cdot p_{+1})}{2} & \text{if } n \cdot p \text{ is integer} \\ n \cdot p_{+1} & \text{if } n \cdot p \text{ is not integer} \end{cases}$$

④ Covariance

$$Cov(x, y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{(n-1)}$$

⑤ Correlation

$$r_{xy} = \frac{Cov(x, y)}{\sigma_x \cdot \sigma_y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

⑥ Point Bi-Serial Correlation

$$r_{pb} = \frac{(\bar{Y}_0 - \bar{Y}_1)}{\sigma_x} \cdot \sqrt{p_0 p_1}$$

where, $\bar{Y}_0, \bar{Y}_1$ = mean of observations (each category)
σ_x = std dev. of all observations
p_0, p_1 = relative frequencies of each category

WEEKS 5-8

① $^{n}P_r = \dfrac{n!}{(n-r)!}$ $\rightarrow$ without repitition $\qquad$ $^{n}P_r = (n)^r$ $\rightarrow$ with repitition

② Circular Permutation
↳ Clockwise $\neq$ anti-clockwise $\Rightarrow$ $(n-1)!$
↳ Clockwise = anti-clockwise $\Rightarrow$ $\dfrac{(n-1)!}{2}$

③ $^{n}C_r \times r! = {}^{n}P_r$; $^{n}C_r = \binom{n}{r} = \dfrac{n!}{(n-r)!\, r!}$

$$^{n}C_r = {}^{n-1}C_{r-1} + {}^{n-1}C_r$$

④ For mutually exclusive events E_i :

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

⑤ $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

⑥ $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

⑦ Independent events A and B :

$$P(A \cap B) = P(A) \times P(B)$$
$$P(A \cap B^c) = P(A) \times P(B^c)$$

⑧ Law of Total Probability :

$$P(E) = P(E \cap F) + P(E \cap F^c)$$
$$= P(F) \times P(E|F) + P(F^c) \times P(E|F^c)$$

↳ $P(E) = \sum_{i=1}^{k} P(E|F_i) \times P(F_i)$

⑨ Baye's Theorem :

$$P(F|E) = \frac{P(F) \times P(E|F)}{P(E)} = \frac{P(F \cap E)}{P(E)}$$
$$= \frac{P(F) \times P(E|F)}{P(F) \times P(E|F) + P(F^c) \times P(E|F^c)}$$

$$P(F_k|E) = \frac{P(F_i) \times P(E|F_i)}{\sum_{i=1}^{k} P(F_i) \times P(E|F_k)}$$

WEEKS 9-12

① $\sum_{i=0}^{\infty} \dfrac{\lambda^i}{i!} = e^{\lambda}$

② Expectations of random discrete variable :

$$E(X) = \sum_{i=1}^{n} x_i P(x = x_i)$$

↳ If a and b are constants

$$E(aX + b) = aE(X) + b$$

③ Variance of random discrete variable :

$$Var(X) = E(X - E(X))^2 = E\left(X^2 + (E(X))^2 - 2E(X)X\right)$$
$$= E(X^2) - (E(X))^2$$

↳ If c is a constant :

$$Var(cX) = c^2 \, Var(X)$$

(4) $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$; $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$

(5) IID bernoulli trials → Binomial Random Variable

↳ $X \sim Bin(n, p)$, where n = no. of trials , p = probability of success

↳ Probability of i successes in n trials :

$$P(X=i) = \binom{n}{i} * (p)^i * (1-p)^{n-i}$$

(6) Expectations of Bin(n,p) ⇒ $E(X) = n \cdot p$

(7) Variance of Bin(n,p) ⇒ $Var(X) = n \cdot p \cdot (1-p)$

(8) Hypergeometric Distribution – without replacement

m = no. of target items n = no. of items (sample size)

x = no. of target items drawn N = Population

$$P(X=x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

(9) Expectation of Hypergeometric $E(X) = \frac{n \cdot m}{N}$

(10) Variance of Hypergeometric $Var(X) = n\left(\frac{m}{N}\right)\left(\frac{N-m}{N}\right)\left(\frac{N-n}{N-1}\right)$

↳ If $N >> n$, then Hypergeometric (N, m, n) is similar to $Bin(n,p)$; $p = \frac{m}{N}$

(11) Poisson Distribution ⇒ fixed interval (time / space) → Poisson (λ)

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

(12) Expectation of Poisson Distribution : $E(X) = \lambda$

(13) Variance of Poisson Distribution : $Var(X) = \lambda$

(14) Uniform continuous Distribution : $X \sim U(a, b)$

$$P(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

(15) Cumulative distribution of uniform continuous distribution :

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b) \\ 1 & x \geq b \end{cases}$$

$$P(X \leq x) = \frac{x-a}{b-a}$$

(16) Expectation of a uniform distribution : $E(X) = \frac{b+a}{2}$

(17) Variance of a uniform distribution : $Var(X) = \frac{(b-a)^2}{12}$

(18) Exponential Distribution :

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$P(x) = \lambda e^{-\lambda x}$$

(19) Cumulative distribution of Exponential :

$$P(X \leq x) = 1 - e^{-\lambda x}$$

(20) Expectation of Exponential Distribution : $E(X^n) = \frac{1}{\lambda}$

(21) Variance of Exponential Distribution : $Var(X^n) = \frac{1}{\lambda^2}$

(22) Geometric Distribution : doing Bernoulli trials until a success :

$$P(G \leq k) = 1 - (1-p)^k$$

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

$$E(X) = \frac{1}{p} \qquad Var(X) = \frac{(1-p)}{p^2}$$