Formula Sheet - Final Exam - STAT 260

- $P(B|A) = \dfrac{P(B \cap A)}{P(A)}$, for $P(A) > 0$.
- $P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_k)P(A_k)$, for $A_1, \ldots, A_k$ mutually exclusive and exhaustive
- $E(X) = \mu = \displaystyle\sum_{all\ x} x f(x)$
- $E(g(X)) = \displaystyle\sum_{all\ x} g(x) f(x)$
- $V(X) = E(X^2) - [E(X)]^2$
- For a continuous random variable X with pdf $f(x)$,

$$\begin{aligned} P(a \leq X \leq b) &= \int_a^b f(x)\, dx \\ E(X) &= \int_{-\infty}^{\infty} x f(x)\, dx \\ E(g(X)) &= \int_{-\infty}^{\infty} g(x) f(x)\, dx \end{aligned}$$

- $E(XY) = \sum_x \sum_y x \cdot y \cdot f(x, y)$
- $Cov(X, Y) = E(XY) - E(X)E(Y)$
- If $X \sim N(\mu, \sigma)$, then: $Z = \frac{X - \mu}{\sigma}$
- If $X \sim Exp(\lambda)$, $F(x) = 1 - e^{-\lambda x}$

Distribution	**Notation**	**pmf/pdf**
Binomial	$X \sim \text{Bin}(n, p)$	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$
Poisson	$X \sim \text{Poisson}(\lambda)$	$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$
Uniform	$X \sim \text{Uniform}(a, b)$	$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b$
Gamma	$X \sim \text{Gamma}(\alpha, \beta)$	$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \quad x > 0$
Exponential	$X \sim \text{Exponential}(\lambda)$	$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$

- **Confidence Interval**: estimate $\pm$ (c.v.)(e.s.e)

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$\bar{x} \pm t_{n-1,\alpha/2} \frac{s}{\sqrt{n}}$$

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- **Sample size** $n = \left\lceil \left(\dfrac{z_{\frac{\alpha}{2}} \cdot \sigma}{d} \right)^2 \right\rceil$ or $n = \left\lceil \left(\dfrac{z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})}}{d} \right)^2 \right\rceil$

- **Test Statistics:** $\dfrac{\text{estimate- true parameter under null}}{\text{estimated standard error}}$

  $Z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

  $Z = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \approx N(0,1)$

  $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$

  $Z = \dfrac{\hat{p} - p}{\sqrt{p(1-p)/n}} \sim N(0,1)$

  $Z = \dfrac{(\bar{x_1} - \bar{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0,1)$

  $t = \dfrac{(\bar{x_1} - \bar{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}$

  $t = \dfrac{(\bar{x_1} - \bar{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_\nu$, where $\nu = \dfrac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$

  $Z = \dfrac{(\hat{p_1} - \hat{p_2}) - (p_1 - p_2)}{\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}} \sim N(0,1)$

  $t = \dfrac{\bar{x}_D - \mu_D}{\frac{s_D}{\sqrt{n_D}}} \sim t_{n_D-1}$, where $\bar{x}_D$ and s_D are the mean and standard deviation of D_i respectively, and $D_i = x_i - y_i$.

- **P-value** (against H_0)
  - *Overwhelming* or *Very Strong* Evidence if p-value ≤ 0.01
  - *Strong* Evidence if $0.01 <$ p-value ≤ 0.05
  - *Weak* or *Moderate* Evidence if $0.05 <$ p-value ≤ 0.10
  - *No Evidence* if $0.10 <$ p-value