FINAL EXAM FORMULA SHEET

5A

Miscellaneous Formulae

Please note- it is up to you to understand what each formula means, and it is also up to you to know which formula you need to use in a given situation. We (the Course Staff) will not be able to answer any questions about these formulas during the Exam.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$s_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

$$s_X = \sqrt{s_X^2}$$

$$\text{IQR} = Q_3 - Q_1$$

$$\text{range}(X) = \max\{X\} - \min\{X\}$$

$$0 \leq \mathbb{P}(A) \leq 1$$

$$\mathbb{P}(\varnothing) = 0$$

$$\mathbb{P}(\Omega) = 1$$

$$\mathbb{P}(A^{\complement}) = 1 - \mathbb{P}(A)$$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)} \text{ provided that } \mathbb{P}(F) \neq 0$$

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(F \mid E) \cdot \mathbb{P}(E)}{\mathbb{P}(F)} \text{ provided that } \mathbb{P}(E) \neq 0 \text{ and } \mathbb{P}(F) \neq 0$$

$$E \perp F \text{ if any of: } \mathbb{P}(E \mid F) = \mathbb{P}(E); \quad \mathbb{P}(F \mid E) = \mathbb{P}(F); \quad \mathbb{P}(E \cap F) = \mathbb{P}(E) \cdot \mathbb{P}(F)$$

$$0! = 1$$

$$\mathbb{P}(E) = \mathbb{P}(E \cap F) + \mathbb{P}(E \cap F^{\complement})$$

$$n! = n \times (n-1) \times \cdots \times 2 \times 1$$

$$(n)_k = \frac{n!}{(n-k)!}$$

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

$$\mathbb{P}(X = k) \geq 0$$

$$\sum_{\text{all } k} \mathbb{P}(X = k) = 1$$

$$\text{SD}(X) = \sqrt{\text{Var}(X)}$$

$$\text{Var}(X) = \sum_{\text{all } k}(k - \mathbb{E}[X])^2 \cdot \mathbb{P}(X = k) = \left(\sum_{\text{all } k} k^2 \cdot \mathbb{P}(X = k)\right) - (\mathbb{E}[X])^2$$

$\mathbb{E}[aX + b + c] = a \cdot \mathbb{E}[X] + b \cdot \mathbb{E}[Y] + c$		$\text{Var}(aX + bY + c) = a^2\text{Var}(X) + b^2\text{Var}(Y)$
$\text{TS} = \dfrac{\widehat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	$\text{TS} = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	$\text{TS} = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$
$\text{TS} = \dfrac{\overline{Y} - \overline{X}}{\sqrt{\frac{s_X^2}{n_1} + \frac{s_Y^2}{n_2}}}$	$\text{TS} = \dfrac{\widehat{\beta}_1}{\text{SD}(\widehat{\beta})_1}$	$\widehat{\beta}_0 = \overline{y} - \widehat{\beta}_1 \cdot \overline{x}$
$r = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(\dfrac{x_i - \overline{x}}{s_X}\right)\left(\dfrac{y_i - \overline{y}}{s_Y}\right)$	$\widehat{\beta}_1 = \dfrac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \dfrac{s_Y}{s_X} \cdot r$	
$\text{df} = \text{round}\left\{\dfrac{\left[\left(\frac{s_X^2}{n_1}\right) + \left(\frac{s_Y^2}{n_2}\right)\right]^2}{\frac{\left(\frac{s_X^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_Y^2}{n_2}\right)^2}{n_2 - 1}}\right\}$		$\text{TS} \overset{H_0}{\sim} t_{\text{df}}$
$\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x$	$e_i = y_i - \widehat{y}_i$	$\text{RSS} = \sum_{i=1}^{n} e_i^2$

Binomial Distribution: $X \sim \text{Bin}(n, p)$

$S_X = \{0, 1, 2, \cdots, n\}$	$\mathbb{E}[X] = np$	$\text{Var}(X) = np(1-p)$
$\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$ if $k \in S_X$ and 0 otherwise		

Uniform: $X \sim \text{Unif}(a,\ b)$

$S_X = [a, b]$	$\mathbb{E}[X] = \dfrac{a+b}{2}$	$\text{Var}(X) = \dfrac{(b-a)^2}{12}$
$f_X(X) = \dfrac{1}{b-a}$ if $x \in S_X$ and 0 otherwise		

Normal: $X \sim \mathcal{N}(\mu,\ \sigma)$

$S_X = \mathbb{R} = (-\infty, \infty)$	$\mathbb{E}[X] = \mu$	$\text{Var}(X) = \sigma^2$
$f_X(X) = \dfrac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{ -\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2 \right\}$		$Z = \left(\dfrac{X-\mu}{\sigma}\right) \sim \mathcal{N}(0,\ 1)$

Linear Combinations of Normally-Distributed Random Variables

If $X \sim \mathcal{N}(\mu_X,\ \sigma_X)$ and $Y \sim \mathcal{N}(\mu_Y,\ \sigma_Y)$ with $X \perp Y$, then

$$(aX + bY + c) \sim \mathcal{N}\left(a\mu_X + b\mu_Y + c\ ,\ \sqrt{a^2\sigma_X^2 + b^2\sigma_Y^2}\right)$$

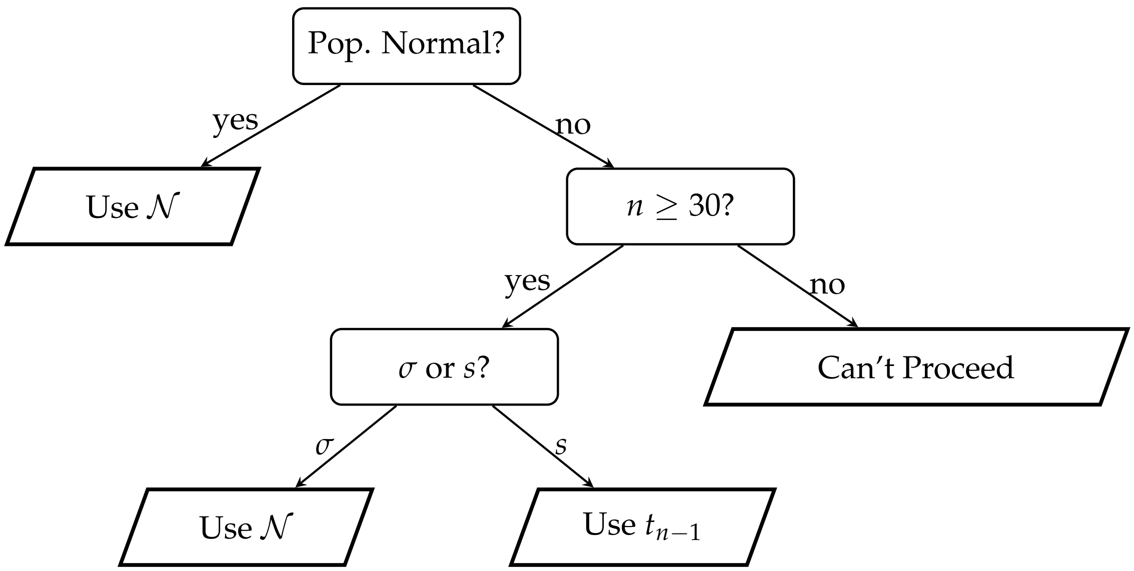
Central Limit Theorem for Proportions

Given a population with proportion p, define $\widehat{P}$ to be the sample proportion. Then

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

provided: **(1)** $np \geq 10$ and **(2)** $n(1-p) \geq 0$ **—OR— (1)** $n\widehat{p} \geq 10$ and **(2)** $n(1-\widehat{p}) \geq 0$

Flowchart for the Sampling Distribution of $\overline{X}$



ANOVA

- $\text{MS}_\text{G} = \dfrac{\text{SS}_\text{G}}{k-1}$
  (Btwn. Groups)
- $\text{MS}_\text{E} = \dfrac{\text{SS}_\text{E}}{n-k}$
  (Chance/noise)
- $F = \dfrac{\text{MS}_\text{G}}{\text{MS}_\text{E}} \overset{H_0}{\sim} F_{k-1,\, n-k}$

Assorted Coding Results

- `.ppf(q, *args)` : point-percent function. Description of arguments:
  - `q`: `array_like`; lower tail probability
  - `*args`: parameters of the distribution

- `.cdf(x, *args)` : cumulative distribution function. Description of arguments:
  - `x`: quantiles
  - `*args`: parameters of the distribution

- `.pdf(x, *args)` : probability density function. Description of arguments:
  - `x`: `array_like`; quantiles
  - `*args`: parameter(s) of the distribution