Stat 201: Formulas for final exam
******Note: You may add notes to the front and back of this formula sheet ******

- Sample mean: $\bar{x} = \frac{\sum x}{n}$; Population mean $\mu = \sum[xP(x)]$
- median: the midpoint of the observations when they are ordered from smallest to largest.
- $z = \frac{\text{observation} - \text{mean}}{\text{std.dev}}$
- Residual $= y - \hat{y}$ where $\hat{y} = a + bx$
- r (correlation coefficient) $= (+\ or\ -)\sqrt{R^2}$
- $P(A) = \frac{N(A)}{N(S)}$ $\quad P(A^c) = 1 - P(A)$ $\quad P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$
- $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$ $\quad P(A \text{ and } B) = P(A|B)P(B)$
- Mutually exclusive (disjoint) events: $P(A \text{ and } B) = 0$
- Independent events: $P(A|B) = P(A)$; $\quad P(B|A) = P(B)$; $\quad P(A \text{ and } B) = P(A)P(B)$
- Properties of a probability distribution: 1. $0 \leq P(x) \leq 1$ $\quad$ 2. $\sum P(x) = 1$
- Binomial probability: $P(X = x) = \frac{n!}{x!(n-x)!}p^x q^{n-x}$ where $x = 0, 1, 2, \ldots, n$
- For Binomial random variable X: $\mu_X = np$ and $\sigma_X = \sqrt{npq}$
- The mean and standard error of a sample proportion is: $\mu_{\hat{p}} = p$ and $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$
- The mean and standard error of a sample mean is: $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
- General way to construct confidence interval: point estimate $\pm$ margin of error:
  - Population proportion: $\hat{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
  - One population mean (σ known): $\bar{x} \pm z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$
  - One population mean (σ unknown): $\bar{x} \pm t_{\frac{\alpha}{2},df}\frac{s}{\sqrt{n}}$ where $df = n - 1$

- Diff. between indep. population proportions: $(\hat{p}_1-\hat{p}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
- Diff. between indep. population means: $(\bar{x}_1-\bar{x}_2) \pm t_{\frac{\alpha}{2},df} \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$, $df = n_1+n_2-2$
- The value of $z_{\frac{\alpha}{2}}$ is determined by following table:

Confidence level $1-\alpha$	$z_{\frac{\alpha}{2}}$
0.90	1.645
0.95	1.96
0.99	2.58

- General 5-step procedure to construct hypothesis testing (p-value approach):

1. State H_0 and H_a
2. Check assumptions
3. Calculate test statistic:
   * Population proportion: $z^* = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
   * Population mean (σ known): $z^* = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$
   * population mean (σ unknown): $t^* = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$
   * Diff. between indep. population proportions: $z^* = \frac{\hat{p}_1-\hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$
   * Diff. between indep. population means: $t^* = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$
4. Calculate p-value by hand / by StatCrunch
5. Make decision and **write thorough interpretation**:
   * p-value $\leq \alpha \Longrightarrow$ Reject H_0
   * p-value $> \alpha \Longrightarrow$ Fail to reject H_0