Review for the third midterm

1. Statistic is a function of sample observation

2. Sample mean
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$
Sample variance
$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

3. Central limit theorem: When n is large, $\bar{x}$ is approximately normal with mean μ and variance $\frac{\sigma^2}{n}$

4. Y binomial n, p, $\mu = np$, $\sigma = \sqrt{np(1-p)}$
   - If $\mu \pm 2\sigma$ in $(0, n)$, Y is approximately normal
   - Continuity correction, e.g.,
   $$P(Y \geq 6) \overset{CC}{=} P(Y \geq 5.5)$$
   - $$Z = \frac{Y - \mu}{\sigma}$$

5. $U = \frac{(n-1)S^2}{\sigma^2}$ has χ^2 distribution with $\nu = n - 1$
$$\mathbb{E}(S^2) = \sigma^2, \qquad V(S^2) = \frac{2\sigma^4}{n-1}$$

6. An estimator $\hat{\theta}$ of θ is unbiased if
$$\mathbb{E}(\hat{\theta}) = \theta$$

7. $[g_1(\hat{\theta}), g_2(\hat{\theta})]$ is a $1 - \alpha$ confidence interval of θ if
$$P\left(g_1(\hat{\theta}) \leq \theta \leq g_2(\hat{\theta})\right) = 1 - \alpha$$

- When n is large,

$$\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

  is a $1-\alpha$ c.i. of μ. Replace σ by s if σ is unknown.

$$n = \left(\frac{z_{\alpha/2}\sigma}{B}\right)^2$$

- Y binomial n, p. Let $\hat{p} = \frac{Y}{n}$ Then

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

  is a $1-\alpha$ c.i. of p.

$$n = \frac{z_{\alpha/2}^2 \hat{p}(1-\hat{p})}{B^2}$$

- If n is not large, σ is unknown and population is normal, then,

$$T = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

  has a t-distribution with $\nu = n-1$. In this case,

$$\bar{x} \pm t_{\alpha/2}(\nu)\frac{s}{\sqrt{n}}$$

  is a $1-\alpha$ c.i. of μ.
  If σ is known, then

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

  is a $1-\alpha$ c.i. of μ.
-

$$\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}\right]$$

  is a $1-\alpha$ c.i. of σ^2

8. If $X_1, X_2, \cdots, X_n$ is a sample from a population with p.f. or p.d.f. $f(x_i, \theta)$, then the MLE $\hat{\theta}$ of θ maximize

$$L(\theta) = f(x_1, \theta) \cdots f(x_n, \theta)$$

   Note: Most of times, we maximize $\ln L(\theta)$