Stat 401 - 12/6/17

Sec 5.3 - continued

Large Sample Inference for Proportions

Let $X_1, \ldots, X_n$ be a random sample from a Bernoulli Dist w/ prob of success p.

$$X_i \overset{iid}{\sim} Ber(p)$$

Let $\hat{p}$ be the sample proportion of successes.

We can show

$$\frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \xrightarrow{D} N(0,1)$$

why? $\hat{p} = \bar{X}$

$$\hat{p} = \frac{\#\text{ successes (sample)}}{n \text{ (sample size)}} = \frac{X_1 + \cdots + X_n}{n}$$

$$= \bar{X}$$

Let $Y = X_1 + \cdots + X_n \sim \text{Bin}(n, p)$

$$\mathbb{E}[\hat{p}] = \mathbb{E}\left[\frac{Y}{n}\right] = \frac{1}{n}\mathbb{E}[Y]$$

$$= \frac{1}{n} \cdot np = p$$

$$\text{Var}(\hat{p}) = \text{Var}\left(\frac{Y}{n}\right) = \frac{1}{n^2}\text{Var}(Y)$$

$$= \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

$$\frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} = \frac{\sqrt{n}\,(\hat{p} - p)}{\sqrt{\hat{p}(1-\hat{p})}}$$

$$= \frac{\sqrt{n}\,(\overset{\bar{X}}{\hat{p}} - \overset{\mu}{p})}{\underset{\sigma/\sqrt{n}}{\sqrt{p(1-p)}}} \cdot \frac{\sqrt{p(1-p)}}{\sqrt{\hat{p}(1-\hat{p})}}$$

$$\frac{\sqrt{n}\,(\hat{p} - p)}{\sqrt{p(1-p)}} \xrightarrow{D} N(0,1) \qquad \frac{\sqrt{p(1-p)}}{\sqrt{\hat{p}(1-\hat{p})}} \xrightarrow{P} 1 \quad \left(\text{Ex } 5.1.2\right)$$

$$\xrightarrow{D} (1)N(0,1) = N(0,1).$$

Large Sample Inference for χ^2-Tests

Suppose $Y_n \sim Bin(n, p)$

By Normal Approx to Binom

$$\frac{Y_n - np}{\sqrt{np(1-p)}} \xrightarrow{D} N(0,1) \quad \text{by CLT}$$

Let $g(X) = X^2$ continuous.

then by properties

$$\left(\frac{Y_n - np}{\sqrt{np(1-p)}}\right)^2 \xrightarrow{D} [N(0,1)]^2$$

$$= Z^2 \sim \chi^2(1)$$

Final Review

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\, dy$$

Old Material

- Find PMF / PDF from CDF.
- Find CDF from PMF / PDF
- Find $\mathbb{E}[g(X)]$
- Find $Var(X)$
- Transformation
- Relationship between covariance and independence.
- Chebyshev's Inequality.
- $\mathbb{E}[\bar{X}] = \mathbb{E}[X] = \mu$
- $Var(\bar{X}) = \dfrac{\sigma^2}{n}$
- MGF's. $M_X(t) = \mathbb{E}[e^{tX}]$.
- $Z^2 \sim \chi^2(1)$

New Material

- Find Z-scores $Z = \dfrac{X-\mu}{\sigma}$
- Find Normal Dist probabilities
- Find value of z_p
  Recall $\Phi(z_p) = p$

- If $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

  then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

  is exactly Normal

- Bivariate Normal

$(X,Y) \sim N_2(\ldots)$ — marginal dist. are Normal

Find $P(X<2)$ — how to convert to z-scores to find prob.

— parameters are $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$

- Mixture Distributions.

$$E[X] = \sum_{i=1}^{k} p_i \mu_i = \bar{\mu}$$

$$Var[X] = \sum_{i=1}^{k} p_i \sigma_i^2 + \sum_{i=1}^{k} p_i (\mu_i - \bar{\mu})^2$$

Review my HW sol. for Sec 3.7!

- Convergence in Probability.

$$\lim_{n \to \infty} \mathbb{P}\{|X_n - X| > \varepsilon\} = 0$$

OR $$\lim_{n \to \infty} \mathbb{P}\{|X_n - X| < \varepsilon\} = 1$$

Properties:

1. If $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ then

   a) $X_n + Y_n \xrightarrow{P} X+Y$

   b) $X_n Y_n \xrightarrow{P} XY$

2. If $X_n \xrightarrow{P} X$ and a = constant then

$$aX_n \xrightarrow{P} aX$$

3. If $X_n \xrightarrow{P} X$ and g is a continuous function then

$$g(X_n) \xrightarrow{P} g(X)$$

- Weak Law of Large #s.

$$X_1, \ldots, X_n \overset{iid}{\sim} (\mu, \sigma^2)$$

where $\sigma^2 < \infty$

Then $\bar{X} \xrightarrow{P} \mu$

- Convergence In Distribution.

If $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$ for all $x \in C(F_X)$ then $X_n \xrightarrow{D} X$

- If $\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$

then $X_n \xrightarrow{D} X$

Properties:

1. $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X$
2. If $X_n \xrightarrow{D} X$ and $g(x)$ is a continuous function then

$$g(X_n) \to g(X)$$

3. If $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} a$

constant

then $X_n Y_n \xrightarrow{D} aX$

$$X_n + Y_n \xrightarrow{D} X + a$$

Degenerate Distribution at c

i.e. $X = c \leftarrow$ const.

$$f_X(x) = \begin{cases} 1, & x = c \\ 0, & \text{o.w.} \end{cases}$$

$$F_X(x) = \begin{cases} 0, & x < c \\ 1, & x \geq c \end{cases}$$

$$M_X(t) = e^{ct}$$

- Central Limit Thm

If $X_1, \dots, X_n \overset{iid}{\sim} (\mu, \sigma^2)$

where $\sigma^2 < \infty$

then

$$\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$$

(this is approximate).

we could also say

$$\bar{X}_n \xrightarrow{D} N\left(\mu, \frac{\sigma^2}{n}\right)$$

ex)

$$P(\_\_ < \bar{X} < \_\_)$$

$$\text{vs} \quad P(\_\_ < X < \_\_)$$