Econ 241A Probability, Statistics and Econometrics
Fall 2017

Problem Set 5

1. (from Goldberger) Let $\bar{X}$ and S^2 denote the sample mean and sample variance in random sampling, sample size 15, from a n$(10, 100)$ population. Find the probability of each of these events:

$$A = \{\bar{X} \leq 14.9\} \qquad B = \{5.1 \leq \bar{X} \leq 14.9\} \qquad C = \{S^2 \leq 92.04\}$$
$$D = B \cap C \qquad E = \{\sqrt{15}(\bar{X} - 10)/S \leq 1.746\} \qquad F = \{\bar{X} \leq 10 + 0.53S\}$$

- We know that $\bar{X} \sim n(10, 100/15)$ so,

$$P(A) = P(\bar{X} \leq 14.9) = P\left(\frac{\bar{X} - 10}{10/\sqrt{15}} \leq \frac{14.9 - 10}{10/\sqrt{15}}\right) = P\left(z \leq 1.8978\right) = 0.9711$$

$$\begin{aligned} P(B) = P(5.1 \leq \bar{X} \leq 14.9) &= P(A) - P\left(z \leq \frac{5.1 - 10}{10/\sqrt{15}}\right) \\ &= 0.9711 - P(z \leq -1.8978) = 0.9711 - 0.0289 = 0.9423 \end{aligned}$$

- $$P(C) = P((n-1)S^2/\sigma^2 \leq (n-1)92.04/\sigma^2) = P((n-1)S^2/\sigma^2 \leq 14 \times /100)$$

  We know that $(n-1)S^2/\sigma^2 \sim \chi^2_{(n-1)}$, then

$$\begin{aligned} P(C) &= P((n-1)S^2/\sigma^2 \leq 14 \times 92.04/100) \\ &= P(\chi^2_{(14)} \leq 12.8856) = 0.4644 \end{aligned}$$

- Since $\bar{X}$ and S^2 are independent then

$$P(D) = P(B \cap C) = P(B)P(C) = 0.4376$$

- $$\begin{aligned} P(E) &= P(\sqrt{15}(\bar{X} - 10)/S \leq 1.746) \\ &= P(t_{n-1} \leq 1.746) \\ &= 0.9486 \end{aligned}$$

- $$\begin{aligned} P(F) &= P(\sqrt{15}(\bar{X} - 10)/S \leq 0.53 \times \sqrt{15}) \\ &= P(t_{14} \leq 1.746) \\ &= 0.9704 \end{aligned}$$

Note: You will need to use tables (or excel) for evaluating the cdf's of the normal distribution, the Student's t distribution and the Chi-squared distribution.

In addition, solve the following problems from Casella and Berger: 5.1, 5.3, 5.5, 5.11 and 5.15.

5.1 If X = number of people that is color blind in a sample of size n, then $X \sim Binomial(n, p)$ with $p = 0.01$ (a "success" here means having a color blind person). We are interested in $P(X > 0) = 1 - P(X = 0) = 1 - (0.99)^n > 0.95$ which happens if and only if $n > log(0.05)/log(0.99) \approx 299$.

5.3 $Y_i \sim Bernoulli(p_i = P(X_i > \mu)) = Bernoulli(1 - F_X(\mu))$ and since $\{X_i\}_{i=1}^n$ is iid then $\{Y_i\}_{i=1}^n$ is iid. Therefore, $\sum_{i=1}^n Y_i \sim Binomial(n, p)$ with $p = 1 - F_X(\mu)$.

5.5 Let $Y = X_1 + X_2 + \cdots + X_n = F_Y(y)$ and $\bar{X} = Y/n$ so $Y = n\bar{X}$. Then by the transformation method,

$$F_{\bar{X}}(a) = F_Y(n\bar{X}) \left| \frac{dY}{d\bar{X}} \right|$$
$$F_{\bar{X}}(a) = F_Y(n\bar{X})n$$

Here $\frac{dY}{d\bar{X}} = n$ is the Jacobian of the transformation.

5.11 It follows from the Jensen inequality applied to $f(x) = x^2$, a convex function in x. The strict inequality follows from excluding the case when $\sigma = 0$ in which case $S = 0$.

5.15 (a)

$$\begin{aligned}
\bar{X}_{n+1} &= \sum_{i=1}^{n+1} \frac{X_i}{n+1} \\
&= \frac{X_{n+1} + \sum_{i=1}^n X_i}{n+1} \\
&= \frac{X_{n+1} + n\bar{X}_n}{n+1}
\end{aligned}$$

(b)

$$\begin{aligned}
nS_{n+1}^2 &= \frac{n}{(n+1)-1} \sum_{i=1}^{n+1} (X_i - \bar{X}_{n+1})^2 \\
&= \sum_{i=1}^{n+1} \left(X_i - \frac{X_{n+1} + n\bar{X}_n}{n+1} \right)^2 \\
&= \sum_{i=1}^{n+1} \left(X_i - \frac{X_{n+1}}{n+1} - \frac{n\bar{X}_n}{n+1} \right)^2 \\
&= \sum_{i=1}^{n+1} \left((X_i - \bar{X}_n) - \left(\frac{X_{n+1}}{n+1} - \frac{\bar{X}_n}{n+1} \right) \right)^2 \\
&= \sum_{i=1}^{n+1} \left[(X_i - \bar{X}_n)^2 - 2(X_i - \bar{X}_n) \left(\frac{X_{n+1}}{n+1} - \frac{\bar{X}_n}{n+1} \right) + \left(\frac{X_{n+1}}{n+1} - \frac{\bar{X}_n}{n+1} \right)^2 \right] \\
&= \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 + (X_{n+1} - \bar{X}_n)^2 - 2\frac{(X_{n+1} - \bar{X}_n)^2}{n+1} + \frac{(X_{n+1} - \bar{X}_n)^2}{(n+1)^2}(n+1) \\
&= (n-1)S_n^2 + \frac{n}{n+1}(X_{n+1} - X_n)^2
\end{aligned}$$

Second equality uses 5.15 (a), forth adds and substracts $\bar{X}_n$ and the sixth that the sum of deviations from the mean are zero.