Solution HW 5

October 10, 2019

Section 7.7 of Rice; Exercises 1, 3, 23, 24, 43, 48, 50, 51 **7.1 -**
Here $x_1 = 1, x_2 = 2, x_3 = 2, x_4 = 4, x_5 = 8$ and $N = 5$. Thus, mean $\mu = \frac{1}{5}\sum_{i=1}^{5} x_i = 3.4$ and variance $\sigma^2 = \frac{1}{5}\sum_{i=1}^{5} x_i^2 - 3.4^2 = 6.24$.

Possible samples of size 2	Sample mean values, $\bar{x}$
$(x_1, x_2) = (1, 2)$	$\frac{3}{2}$
$(x_1, x_3) = (1, 2)$	$\frac{3}{2}$
$(x_1, x_4) = (1, 4)$	$\frac{5}{2}$
$(x_1, x_5) = (1, 8)$	$\frac{9}{2}$
$(x_2, x_3) = (2, 2)$	2
$(x_2, x_4) = (2, 4)$	3
$(x_2, x_5) = (2, 8)$	5
$(x_3, x_4) = (2, 4)$	3
$(x_3, x_5) = (2, 8)$	5
$(x_4, x_5) = (4, 8)$	6

The sampling distribution of $\bar{X}$ can be described by it's PMF

$$P_{\overline{X}}(\bar{x}) = P(\overline{X} = \bar{x}) = \begin{cases} \dfrac{1}{10} & \text{if } \bar{x} = 2, 2.5, 4.5, 6 \\ \dfrac{2}{10} & \text{if } \bar{x} = 1.5, 3, 5 \\ 0 & \text{otherwise} \end{cases}$$

Now treat $\bar{X}$ as any discrete RV, then after standard expected value calculations, we have $E(\bar{X}) = 3.4$ and $Var(\bar{X}) = 2.34$.
Theorem A says, $E(\bar{X}) = \mu = 3.4$ and Theorem B says, $var(\bar{X}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{6.24}{2}\left(\frac{5-2}{5-1}\right) = 2.34$.

7.3 -
Which of the following is a random variable?

a. The population mean
b. The population size, N
c. The sample size, n

d. The sample mean
e. The variance of the sample mean
f. The largest value in the sample
g. The population variance
h. The estimated variance of the sample mean

Answer. d, f, h.

7.23 (a) The standard error of an estimated proportion is

$$\sqrt{\frac{p(1-p)}{n}(1-\frac{n-1}{N-1})}.$$

It's maximized when $p = 1/2$.

(b) According to Corollary B in Section 7.3.2, an unbiased estimate of $Var(\hat{p})$ is

$$s_{\hat{p}}^2 = \frac{\hat{p}(1-\hat{p})}{n-1}\Big(1-\frac{n}{N}\Big)$$

.

Thus

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}\Big(1-\frac{n}{N}\Big)} \leq \frac{1}{2}\sqrt{\frac{1}{n-1}\Big(1-\frac{n}{N}\Big)} = \frac{1}{2}\sqrt{\frac{N-n}{N(n-1)}}$$

(c) According to the central limit theory,

$$\hat{p} \sim N(p, s_{\hat{p}})$$

In this problem, we need to prove

$$P\Big(\hat{p} - \sqrt{\frac{N-n}{N(n-1)}} < p < \hat{p} + \sqrt{\frac{N-n}{N(n-1)}}\Big) \geq 0.95.$$

The above probability is equivalent to

$$P(p - \sqrt{\frac{N-n}{N(n-1)}}) < \hat{p} < p + \sqrt{\frac{N-n}{N(n-1)}})$$

$$= P\Big(\frac{-\sqrt{\frac{N-n}{N(n-1)}}}{s_{\hat{p}}} < \frac{\hat{p}-p}{s_{\hat{p}}} < \frac{\sqrt{\frac{N-n}{N(n-1)}}}{s_{\hat{p}}}\Big) = \Phi\Big(\frac{\sqrt{\frac{N-n}{N(n-1)}}}{s_{\hat{p}}}\Big) - \Phi\Big(-\frac{\sqrt{\frac{N-n}{N(n-1)}}}{s_{\hat{p}}}\Big).$$

According to (b), $s_{\hat{p}} \leq \frac{1}{2}\sqrt{\frac{N-n}{N(n-1)}}$, thus

$$\Phi\Big(\frac{\sqrt{\frac{N-n}{N(n-1)}}}{s_{\hat{p}}}\Big) - \Phi\Big(-\frac{\sqrt{\frac{N-n}{N(n-1)}}}{s_{\hat{p}}}\Big) \geq \Phi\Big(\frac{\sqrt{\frac{N-n}{N(n-1)}}}{\frac{1}{2}\sqrt{\frac{N-n}{N(n-1)}}}\Big) - \Phi\Big(-\frac{\sqrt{\frac{N-n}{N(n-1)}}}{\frac{1}{2}\sqrt{\frac{N-n}{N(n-1)}}}\Big)$$

$$= \Phi(2) - \Phi(-2) = 0.95$$

7.24 (a)

$$E(\bar{X}_c) = \sum_{i=1}^{n} c_i u = u \sum_{i=1}^{n} c_i$$

Thus if $\sum_{i=1}^{n} c_i = 1$, the estimate is unbiased.

(b)

$$Var(\bar{X}_c) = \sum_{i=1}^{n} c_i^2 \sigma^2.$$

Let

$$G(c, \lambda) = \sum_{i=1}^{n} c_i^2 \sigma^2 + \lambda(\sum_{i=1}^{n} c_i - 1)$$

In order to minimize $G(c, \lambda)$, let $\frac{\partial G(c,\lambda)}{\partial c_i} = 0$ and $\frac{\partial G(c,\lambda)}{\partial \lambda} = 0$. Thus we have

$$\frac{\partial G(c, \lambda)}{\partial c_i} = 0 \Rightarrow 2c_i\sigma^2 + \lambda = 0 \Rightarrow c_i = \frac{-\lambda}{2\sigma^2}$$

and

$$\frac{\partial G(c, \lambda)}{\partial \lambda} = 0 \Rightarrow \sum_{i=1}^{n} c_i = 0,$$

which lead to $\lambda = -\frac{2\sigma^2}{n}$ and $c_i = \frac{1}{n}$

7.43 $R = \frac{\bar{V}}{\bar{O}} = 1.6/2.2 = 0.73.$

$$\begin{aligned} s_R^2 &= \frac{1}{n}\Big(1 - \frac{n-1}{N-1}\Big)\frac{1}{\bar{O}^2}(R^2 s_o^2 + s_v^2 - 2Rs_{ov}) \\ &= \frac{1}{100}\Big(1 - \frac{100-1}{8000-1}\Big)\frac{1}{2.2^2} \times (0.73^2 \times 0.7^2 + 0.8^2 - 2 \times 0.73 \times 0.7 \times 0.8 \times 0.85) = 0.0004 \end{aligned}$$

Thus $S_R = \sqrt{s_R^2} = 0.02$

$$E(R) = 0.73 + \frac{1}{100}(1 - 99/7999)\frac{1}{2.2^2}(0.73 \times 0.6^2 - 0.85 \times 0.7 \times 0.8) = 0.73$$

According to central limit theory,

$$R \sim N(0.73, 0.02).$$

Thus a 95 percent confidence interval for R is

$$0.73 \pm 0.02 * 1.96 = 0.73 \pm 0.04$$

7.48 (a)

$$r = 10000/320 = 31.25$$

(b)In order to find the confidence interval, we need to calculate s_R. Firstly,

$$s_x^2 = \frac{1}{n-1}(\sum X_i^2 - \frac{(\sum X_i)^2}{n}) = \frac{1}{99}(1250 - 320^2/100) = 2.28$$

$$s_y^2 = \frac{1}{n-1}(\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}) = \frac{1}{99}(1100000 - 10000^2/100) = 1010.10$$

$$\begin{aligned} s_{xy} &= \frac{1}{n-1}\sum(X_i - \bar{X})(Y_i - \bar{Y}) \\ &= \frac{1}{n-1}(\sum X_iY_i - \frac{\sum X_i \sum Y_i}{n}) \\ &= \frac{1}{99}(36000 - 320 \times 10000/100) = 40.40 \end{aligned}$$

Finally,

$$s_R^2 = \frac{1}{100}\Big(1 - \frac{100-1}{100000-1}\Big)\frac{1}{3.2^2} \times (31.25^2 \times 2.28 + 1010.10 - 2 \times 31.25 \times 40.40) = 0.694$$

Thus $s_R = \sqrt{s_R^2} = 0.83$

The corresponding 95 percent confidence interval is

$$31.25 \pm 0.83 \times 1.96 = 31.25 \pm 1.63$$

(c) Let $T = N\bar{Y}$,

$$T = N \times \frac{\sum Y_i}{n} = 100000 \times 10000/100 = 10^7$$

and

$$s_T^2 = N^2 s_{\bar{Y}}^2 = 10^{10} \times \frac{s_y^2}{n} \times (1 - \frac{n}{N}) = 10^{10} \times \frac{1100000 - 10000^2/100}{100} \times (1 - 100/100000) = 9.99 \times 10^1 2.$$

Thus the corresponding 90 percent confidence interval is

$$10^7 \pm 1.64 \times \sqrt{9.99 \times 10^1 6} = 10^7 \pm 5.2 \times 10^6$$

7.50 (a)

$$\begin{aligned}
Cov(R,\bar{X}) &= E(\bar{Y}) - E(R)E(\bar{X}) \\
\Rightarrow \frac{Cov(R,\bar{X})}{u_x} &= r - E(R)
\end{aligned} \tag{1}$$

$$\begin{aligned}
(1) \Rightarrow -\frac{\sigma_R \sigma_{\bar{X}} \rho(R,\bar{X})}{u_x} &= E(R) - r \\
\Rightarrow \frac{E(R)-r}{\sigma_R} &= -\frac{\sigma_{\bar{X}} \rho(R,\bar{X})}{u_x} \\
\text{Since } -1 \leq \rho(R,\bar{X}) \leq 1 & \\
\Rightarrow \frac{E(R)-r}{\sigma_R} &\leq \frac{\sigma_{\bar{X}}}{u_x}
\end{aligned}$$

Similarly,

$$\begin{aligned}
(1) \Rightarrow \frac{\sigma_R \sigma_{\bar{X}} \rho(R,\bar{X})}{u_x} &= -(E(R) - r) \\
\Rightarrow -\frac{(E(R)-r)}{\sigma_R} &\leq \frac{\sigma_{\bar{X}}}{u_x}
\end{aligned}$$

Thus

$$\frac{|E(R)-r|}{\sigma_R} \leq \frac{\sigma_{\bar{X}}}{u_x}$$

(b) The upper bound in Problem 43 is

$$0.7/2.2 \times \sqrt{\frac{1}{100} \times (1 - \frac{99}{7999})} = 0.032$$

7.51

$$\begin{aligned}
E(\hat{\theta}_j) &= \frac{1}{p}\sum_{j=1}^{p} E(V_j) \\
&= \frac{1}{p}\sum_{j=1}^{p} pE(\hat{\theta}) - (p-1)E(\hat{\theta}_j) \\
&= \frac{1}{p}\sum_{j=1}^{p} [(p\theta + \frac{pb_1}{n} + \frac{pb_2}{n^2} + \dots) - ((p-1)\theta + \frac{b_1}{m} + \frac{b_2}{m^2(p-1)} + \dots)] \\
&= \theta + (\frac{p}{n} - \frac{1}{m})b_1 + (\frac{p}{n^2} - \frac{1}{m^2(p-1)})b_2 + \dots \\
&= \theta + (\frac{1}{m^2 p} - \frac{1}{m^2(p-1)})b_2 + (\frac{1}{m^3 p^2} - \frac{1}{m^3(p-1)^2})b_3 + \dots
\end{aligned}$$