WS08 - Central Limit Theorem and Method of Moments

Shawn Kim

***Directions**: Please upload a PDF to Gradescope that includes both your written responses and corresponding R code inputs/outputs (if requested) for each problem.*

Special Directions In this worksheet we will use Delta Method to analyze estimators. You will use the well-know thin lens formula to create an estimator. Other useful topics to review related to this worksheet are the Law of Large Numbers & Central Limit Theorem. Be sure to demonstrate the correct use of mathematical notation/reasoning in your work. Clearly show your work/reasoning by entering all necessary algebra/calculations as text or inserting a clear well-cropped image of your work using an R chunk.

Problem 1. The focal length f of an optical instrument is needed. This is determined by using the thin lens formula,

$$\frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{f},$$

where r_1 is the distance from the lens to the actual object and r_2 is the distance from the lens to the image of the object. The distance r_1 is independently measured 36 times and r_2 is independently measured 40 times. The mean of the measurements is the actual distances, 10 centimeters and 18 centimeters, respectively. The standard deviation of the measurement is 0.1 centimeters for r_1 and 0.5 centimeter for r_2.

Problem 1 Part a) Let $\bar{R}_1$ be the distribution of sample means of the 36 measurements from the lens to the object. Use the Law of Large Numbers to determine the mean $\mu_{\overline{R}_1}$ and standard deviation $\sigma_{\overline{R}_1}$ of these measurements from the lens to the object.

$\mu_{\overline{R}_1} = 10$

$\sigma_{\overline{R}_1} = 0.1/\sqrt{36}$

Problem 1 Part b) Let $\bar{R}_2$ be the distribution of sample means of the 40 measurements to the image. We can use the Law of Large Numbers (as in part a) to determine that $\mu_{\overline{R}_2} = 18$ cm and $\sigma_{\overline{R}_2} = 0.5/\sqrt{40}$ cm. Using the Central Limit Theorem, estimate $P\{\bar{R}_2 < 17.9\text{cm}\}$, clearly showing your computations using both the a calculator AND the Table of Probabilities for the standard normal distribution.

Hint: For example, in `R`, we can determine $P(\overline{R}_2 < 17.9)$ using the command `pnorm()`, as seen in the `R` chunk below.

```
# NOTE: The Central Limit Thm. states that bar_R2 will be approx. normally
# distributed NOTE: So, we use pnorm() to determine P(bar_R2 < 17.9) using the
# mean, sd of bar_R2
pnorm(17.9, mean = 18, sd = 0.5/sqrt(40))
```

```
## [1] 0.1029516
```

```
# NOTE: Or we can standardize first and use
pnorm((17.9 - 18)/(0.5/sqrt(40)), mean = 0, sd = 1)
```

```
## [1] 0.1029516
```

Calculator Input (i.e. calculator command): $(17.9 - 18)/(0.5/\sqrt{40})$

Calculator Output: Z = -1.2649

Table Input (i.e. the two values in the vertical/horizontal headers): -1.2/-0.06

Table Output: $P\{\bar{R}_2 < 17.9\text{cm}\}$ is about 0.10383

Problem 1 Part c) For measurements $r_{1,1}, \ldots, r_{1,36}$ and $r_{2,1}, \ldots, r_{2,40}$, estimate the focal length using

$$\frac{1}{\bar{r}_1} + \frac{1}{\bar{r}_2} = \frac{1}{\hat{f}}.$$

Develop an estimator $\hat{f}$ for the focal length f.

Hint: To develop the estimator, you will need to rearrange the given equation by solving for $\hat{f}$.
(Be clear in your algebra.)

$$\frac{1}{\bar{r}_1} + \frac{1}{\bar{r}_2} = \frac{1}{\hat{f}}$$

$$\hat{f} = \frac{1}{\left(\frac{1}{\bar{r}_1} + \frac{1}{\bar{r}_2}\right)}$$

$$\hat{f} = \frac{1}{\frac{\bar{r}_2 + \bar{r}_1}{\bar{r}_1 \bar{r}_2}}$$

$$\hat{f} = \frac{\bar{r}_1 \bar{r}_2}{\bar{r}_2 + \bar{r}_1}$$

Problem 1 Part d) Now, use the Delta Method to analyze the estimator $\hat{f}$ (from part c) by estimating the mean $\mu_{\hat{f}}$ and standard deviation $\sigma_{\hat{f}}$ of the estimator $\hat{f}$.

Hint: You will need to use the multi-variable case for Delta Method (similar to Group Assignment 14). Be very careful to show your work clearly and using proper notation.

Hint: For comparison, here is a simulation of the experimental protocol 10000 times using the `rnorm` command to simulate sample means for r_1 and r_2. This may also be helpful in checking that you derived the correct estimator in part (c). Also, your answers for part (d) should be somewhat similar. Of course, you will need to take the square root to determine the standard deviation from the variance calculation.

```r
set.seed(2021)
r1_bars <- rnorm(10000, 10, 0.1/sqrt(36))
r2_bars <- rnorm(10000, 18, 0.5/sqrt(40))
f_ests <- (r1_bars * r2_bars)/(r1_bars + r2_bars)
mean(f_ests)
```

```
## [1] 6.428701
```

```r
sd(f_ests)
```

```
## [1] 0.01227092
```

$$\mu_{\hat{f}} = \frac{\bar{r}_1 \bar{r}_2}{\bar{r}_1 + \bar{r}_2}$$

$$\mu_{\hat{f}} = \frac{10(18)}{10+18}$$

$$\boxed{\mu_{\hat{f}} = 6.42857}$$

$$\sigma^2_{\hat{f}(\bar{r}_1, \bar{r}_2)} = \left(\frac{d}{dr_1}\left(\frac{r_1 r_2}{r_1 + r_2}\right)\right)^2 \frac{\sigma_{r_1}^2}{n_{r_1}} + \left(\frac{d}{dr_2}\left(\frac{r_1 r_2}{r_1 + r_2}\right)\right)^2 \frac{\sigma_{r_2}^2}{n_{r_2}}$$

$$= \left(\frac{(r_2)^2}{(r_1 + r_2)^2}\right)^2 \frac{\sigma_{r_1}^2}{n_{r_1}} + \left(\frac{(r_1)^2}{(r_1 + r_2)^2}\right)^2 \frac{\sigma_{r_2}^2}{n_{r_2}}$$

$$\sigma_{\hat{f}(\bar{r}_1, \bar{r}_2)} = \sqrt{\left(\frac{18^2}{28^2}\right)^2 \frac{0.1^2}{36} + \left(\frac{10^2}{(28)^2}\right)^2 \frac{0.5^2}{40}}$$

$$\sigma_{\hat{f}(\bar{r}_1, \bar{r}_2)} = \boxed{0.01221}$$

Problem 2. Assume that X, the value on an unfair die, has the following probability mass function.

x	1	2	3	4	5	6
$f_X(x\|\alpha)$	$\frac{1}{6}(1-3\alpha)$	$\frac{1}{6}(1-2\alpha)$	$\frac{1}{6}(1-\alpha)$	$\frac{1}{6}(1+\alpha)$	$\frac{1}{6}(1+2\alpha)$	$\frac{1}{6}(1+3\alpha)$

Recall: During Group Assignment 15, we already determined the following:
Group Assignment 15 Part a) What values for α are possible?
Since the values of $f_X(x|\alpha)$ have to be non-negative, then $-1/3 < \alpha < 1/3$.
Additionally, note that $\sum_{i=1}^{6} f_X(x_i|\alpha) = 1$.
Group Assignment 15 Part b) Find the mean $\mu_X = E_\alpha[X]$ and the variance $\sigma_X^2 = Var_\alpha(X)$ of a single die roll.
$E[X] = \cdots = \frac{7}{2} + \frac{11}{3}\alpha.$
$E[X^2] = \sum_{i=1}^{6} x_i^2 \cdot f_X(x_i|\alpha) = \cdots = \frac{91}{6} + \frac{77}{3}\alpha.$ *So* $var(X) = E[X^2] - (E[X])^2 = \cdots = \frac{35}{12} - \frac{121}{9}\alpha^2$
There is no need to repeat the calculations from class if you fully understand your group's solutions.

Problem 2 Part a) Let $\bar{X}$ be the average of 100 dice rolls. Find the mean $\mu_{\bar{X}} = E_\alpha[\bar{X}]$ and the variance $\sigma_{\bar{X}}^2 = \mathrm{Var}_\alpha(\bar{X})$.

Hint: You will use the Group Assignment solutions and the Law of Large Numbers to compute the mean $\mu_{\bar{X}}$ and the variance $\sigma_{\bar{X}}^2$ of the 100 repeated dice rolls.

Problem 2 Part b) For repeated independent observations, $\mathbf{x} = (x_1, x_2, \ldots, x_{100})$, of this random variable using the same unfair die, determine the Method of Moments estimate $\hat{\alpha}$, that will estimate the value of α related to the unfair die used in the repeated observations.

Hint: You will need to use (and clearly label) the four steps for the Method of Moments procedure for one parameter.

2a) $\mu_{\bar{X}} = E_d[\bar{X}] = \boxed{\frac{7}{2} + \frac{11}{3} d}$

$$\sigma^2_{\bar{X}} = \frac{\frac{35}{12} - \frac{121}{9} d^2}{100} = \boxed{\frac{35}{12(100)} - \frac{121}{9(100)} d^2}$$

b) $E_d[\bar{X}] = \frac{7}{2} + \frac{11}{3}\hat{d} = \bar{X}$ 1) find EX

$\left(\frac{3}{11}\right)\left(\bar{X} - \frac{7}{2}\right) = \cancel{\frac{11}{3}}\hat{d}\left(\cancel{\frac{3}{11}}\right)$ 2) solve for estimator

$$\hat{d} = \left(\frac{3}{11}\right)\bar{X} - \frac{21}{22}$$

$$\boxed{\hat{d} = \left(\frac{3}{11}\right)\bar{X} - \frac{21}{22}}$$