

STAT 3202: Practice 01

Autumn 2018, OSU

Exercise 1

Consider independent random variables X_1 , X_2 , and X_3 with

- $E[X_1] = 1$, $\text{Var}[X_1] = 4$
- $E[X_2] = 2$, $\text{SD}[X_2] = 3$
- $E[X_3] = 3$, $\text{SD}[X_3] = 5$

(a) Calculate $E[5X_1 + 2]$.

Solution:

$$E[5X_1 + 2] = 5 \cdot E[X_1] + 2 = \boxed{7}$$

(b) Calculate $E[4X_1 + 2X_2 - 6X_3]$.

Solution:

$$\begin{aligned} E[4X_1 + 2X_2 - 6X_3] &= 4 \cdot E[X_1] + 2 \cdot E[X_2] - 6 \cdot E[X_3] \\ &= 4 \cdot (1) + 2 \cdot (2) - 6 \cdot (3) = \boxed{-10} \end{aligned}$$

(c) Calculate $\text{Var}[5X_1 + 2]$.

Solution:

$$\text{Var}[5X_1 + 2] = 5^2 \cdot \text{Var}[X_1] = 25 \cdot (4) = \boxed{100}$$

(d) Calculate $\text{Var}[4X_1 + 2X_2 - 6X_3]$.

Solution:

$$\begin{aligned} \text{Var}[4X_1 + 2X_2 - 6X_3] &= 4^2 \cdot \text{Var}[X_1] + 2^2 \cdot \text{Var}[X_2] + (-6)^2 \cdot \text{Var}[X_3] \\ &= 4^2 \cdot (4) + 2^2 \cdot (3^2) + (-6)^2 \cdot (5^2) = \boxed{1000} \end{aligned}$$

Exercise 2

Consider random variables H and Q with

- $E[H] = 3$, $\text{Var}[H] = 16$
- $\text{SD}[Q] = 4$, $E\left[\frac{Q^2}{5}\right] = 3.2$

(a) Calculate $E[5H^2 - 10]$.

Solution:

$$\begin{aligned}\text{Var}[H] &= E[H^2] - (E[H])^2 \\ 16 &= E[H^2] - 3^2 E[H^2] = 25\end{aligned}$$

$$E[5H^2 - 10] = 5 \cdot E[H^2] - 10 = 5 \cdot 25 - 10 = \boxed{115}$$

(b) Calculate $E[Q]$.

Solution:

$$E[Q^2] = 16$$

$$\text{Var}[Q] = (\text{Var}[Q])^2 = 16$$

$$\begin{aligned}\text{Var}[Q] &= E[Q^2] - (E[Q])^2 \\ 16 &= 16 - E[Q]^2\end{aligned}$$

$$E[Q] = 0$$

Exercise 3

Consider a random variable S with probability density function

$$f(s) = \frac{1}{9000}(2s + 10), \quad 40 \leq s \leq 100.$$

(a) Calculate $E[S]$.

Solution:

$$E[S] = \int_{40}^{100} s \cdot \frac{2s + 10}{9000} ds = \boxed{74}$$

(b) Calculate $\text{SD}[S]$.

Solution:

$$E[S^2] = \int_{40}^{100} s^2 \cdot \frac{2s + 10}{9000} ds = 5760$$

$$\text{Var}[S] = E[S^2] - (E[S])^2 = 5760 - 74^2 = 284$$

$$\text{SD}[S] = \sqrt{\text{Var}[S]} = \sqrt{284} \approx \boxed{16.8523}$$

Exercise 4

Consider independent random variables X and Y with

- $X \sim N(\mu_X = 2, \sigma_X^2 = 9)$
- $Y \sim N(\mu_Y = 5, \sigma_Y^2 = 4)$

(a) Calculate $P[X > 5]$.

Solution:

$$P[X > 5] = P\left[\frac{X - \mu_X}{\sigma_X} > \frac{5 - 2}{3}\right] = P[Z > 1] = 1 - P[Z < 1] = 1 - 0.8413 = \boxed{0.1587}$$

```
pnorm(q = 5, mean = 2, sd = sqrt(9), lower.tail = FALSE)
```

```
## [1] 0.1586553
```

```
1 - pnorm(1)
```

```
## [1] 0.1586553
```

(b) Calculate $P[X + 2Y > 5]$.

Solution:

$$E[X + 2Y] = 1 \cdot E[X] + 2 \cdot E[Y] = 12$$

$$\text{Var}[X + 2Y] = 1^2 \cdot \text{Var}[X] + 2^2 \cdot \text{Var}[Y] = 25$$

$$X + 2Y \sim N(\mu_{X+2Y} = 12, \sigma_{X+2Y}^2 = 25)$$

$$P[X + 2Y] = P\left[\frac{X + 2Y - \mu_{X+2Y}}{\sigma_{X+2Y}} > \frac{5 - 12}{5}\right] = P[Z > -1.4] = 1 - P[Z < -1.4] = 1 - 0.0808 = \boxed{0.9192}$$

```
pnorm(q = 5, mean = 12, sd = sqrt(25), lower.tail = FALSE)
```

```
## [1] 0.9192433
```

```
1 - pnorm(-1.4)
```

```
## [1] 0.9192433
```

Exercise 5

Consider random variables Y_1 , Y_2 , and Y_3 with

- $E[Y_1] = 1, E[Y_2] = -2, E[Y_3] = 3$
- $\text{Var}[Y_1] = 4, \text{Var}[Y_2] = 6, \text{Var}[Y_3] = 8$
- $\text{Cov}[Y_1, Y_2] = 1, \text{Cov}[Y_1, Y_3] = -1, \text{Cov}[Y_2, Y_3] = 0$

(a) Calculate $\text{Var}[3Y_1 - 2Y_2]$.

Solution:

$$\text{Var}[3Y_1 - 2Y_2] = 3^2 \cdot \text{Var}[Y_1] + (-2)^2 \cdot \text{Var}[Y_2] + 2(3)(-2) \cdot \text{Cov}[Y_1, Y_2] = \boxed{48}$$

(b) Calculate $\text{Var}[3Y_1 - 4Y_2 + 2Y_3]$.

Solution:

$$\begin{aligned} \text{Var}[3Y_1 - 4Y_2 + 2Y_3] &= (3)^2 \cdot \text{Var}[Y_1] + (-4)^2 \cdot \text{Var}[Y_2] + (2)^2 \cdot \text{Var}[Y_3] \\ &\quad + (2)(3)(-4) \cdot \text{Cov}[Y_1, Y_2] + (2)(3)(2) \cdot \text{Cov}[Y_1, Y_3] + (2)(-4)(2) \cdot \text{Cov}[Y_2, Y_3] \\ &= \boxed{128} \end{aligned}$$

Exercise 6

Consider using $\hat{\xi}$ to estimate ξ .

(a) If $\text{Bias}[\hat{\xi}] = 5$ and $\text{Var}[\hat{\xi}] = 4$, calculate $\text{MSE}[\hat{\xi}]$

Solution:

$$\text{MSE}[\hat{\xi}] = (\text{bias}[\hat{\xi}])^2 + \text{var}[\hat{\xi}] = 5^2 + 4 = \boxed{29}$$

(b) If $\hat{\xi}$ is unbiased, $\xi = 6$, and $\text{MSE}[\hat{\xi}] = 30$, calculate $\text{E}[\hat{\xi}^2]$

Solution:

First we use the definition of MSE to recover the variance of $\hat{\xi}$.

$$\begin{aligned} \text{MSE}[\hat{\xi}] &= (\text{bias}[\hat{\xi}])^2 + \text{var}[\hat{\xi}] \\ 30 &= 0 + \text{var}[\hat{\xi}] \\ \text{var}[\hat{\xi}] &= 30 \end{aligned}$$

Then using the fact that the estimator is unbiased, we have

$$\text{E}[\hat{\xi}^2] = \text{var}[\hat{\xi}] + (\text{E}[\hat{\xi}])^2 = 30 + 6^2 = \boxed{66}$$

Exercise 7

Using the identity

$$(\hat{\theta} - \theta) = (\hat{\theta} - \text{E}[\hat{\theta}]) + (\text{E}[\hat{\theta}] - \theta) = (\hat{\theta} - \text{E}[\hat{\theta}]) + \text{Bias}[\hat{\theta}]$$

show that

$$\text{MSE}[\hat{\theta}] = \text{E} [(\hat{\theta} - \theta)^2] = \text{Var}[\hat{\theta}] + (\text{Bias}[\hat{\theta}])^2$$

Solution:

We compute:

$$\begin{aligned} \text{MSE}[\hat{\theta}] &= \text{E} [(\hat{\theta} - \theta)^2] \\ &= \text{E} \left[(\hat{\theta} - \text{E}[\hat{\theta}] + \text{E}[\hat{\theta}] - \theta)^2 \right] \quad (\text{ using the given identity}) \\ &= \text{E} \left[(\hat{\theta} - \text{E}[\hat{\theta}])^2 + 2 \cdot (\hat{\theta} - \text{E}[\hat{\theta}]) (\text{E}[\hat{\theta}] - \theta) + (\text{E}[\hat{\theta}] - \theta)^2 \right] \quad (\text{ algebra}) \\ &= \text{E} \left[(\hat{\theta} - \text{E}[\hat{\theta}])^2 \right] + 2 \cdot (\text{E}[\hat{\theta}] - \theta) (\text{E}[\hat{\theta} - \text{E}[\hat{\theta}]] + (\text{E}[\hat{\theta}] - \theta)^2) \quad (\text{ linearity of expectation}) \end{aligned}$$

where $(\text{E}[\hat{\theta}] - \theta)$ in the last line is a constant.

Note that:

- $\text{E} \left[(\hat{\theta} - \text{E}[\hat{\theta}])^2 \right] = \text{Var}(\hat{\theta})$, by the definition of variance
- $\text{E} [\hat{\theta} - \text{E}[\hat{\theta}]] = \text{E}[\hat{\theta}] - \text{E}[\hat{\theta}] = 0$
- $(\text{E}[\hat{\theta}] - \theta)^2 = \text{Bias}[\hat{\theta}]^2$, by definition of bias

Substituting these into the last expression above, we have

$$\text{MSE}[\hat{\theta}] = \text{Var}[\hat{\theta}] + (\text{Bias}[\hat{\theta}])^2$$

Exercise 8

Let X_1, X_2, \dots, X_n denote a random sample from a population with mean μ and variance σ^2 .

Consider three estimators of μ :

$$\hat{\mu}_1 = \frac{X_1 + X_2 + X_3}{3}, \quad \hat{\mu}_2 = \frac{X_1}{4} + \frac{X_2 + \dots + X_{n-1}}{2(n-2)} + \frac{X_n}{4}, \quad \hat{\mu}_3 = \bar{X},$$

Calculate the mean squared error for each estimator. (It will be useful to first calculate their bias and variances.)

Solution:

We first calculate the expected value of each estimator.

$$\begin{aligned} \text{E}[\hat{\mu}_1] &= \text{E} \left[\frac{X_1 + X_2 + X_3}{3} \right] = \frac{1}{3}(\mu + \mu + \mu) = \mu \\ \text{E}[\hat{\mu}_2] &= \text{E} \left[\frac{X_1}{4} + \frac{X_2 + \dots + X_{n-1}}{2(n-2)} + \frac{X_n}{4} \right] = \frac{\mu}{4} + \frac{(n-2)\mu}{2(n-2)} + \frac{\mu}{4} = \frac{\mu}{4} + \frac{\mu}{2} + \frac{\mu}{4} = \mu \\ \text{E}[\hat{\mu}_3] &= \text{E}[\bar{X}] = \mu \end{aligned}$$

Thus we see that each estimator is unbiased.

Next we calculate the variance of each estimator.

$$\begin{aligned}\text{Var} [\hat{\mu}_1] &= \text{Var} \left[\frac{X_1 + X_2 + X_3}{3} \right] = \frac{1}{9} (\sigma^2 + \sigma^2 + \sigma^2) = \frac{\sigma^2}{3} \\ \text{Var} [\hat{\mu}_2] &= \text{Var} \left[\frac{X_1}{4} + \frac{X_2 + \dots + X_{n-1}}{2(n-2)} + \frac{X_n}{4} \right] = \frac{\sigma^2}{16} + \frac{(n-2)\sigma^2}{4(n-2)^2} + \frac{\sigma^2}{16} = \frac{n\sigma^2}{8(n-2)} \\ \text{Var} [\hat{\mu}_3] &= \text{Var} [\bar{X}] = \frac{\sigma^2}{n}\end{aligned}$$

Since each estimator is unbiased, and for any estimator

$$\text{MSE} [\hat{\theta}] = \text{E} [(\hat{\theta} - \theta)^2] = \left(\text{bias} [\hat{\theta}] \right)^2 + \text{var} [\hat{\theta}]$$

we see that the resulting mean squared errors are simply the variances.

$$\begin{aligned}\text{MSE} [\hat{\mu}_1] &= \boxed{\frac{\sigma^2}{3}} \\ \text{MSE} [\hat{\mu}_2] &= \boxed{\frac{n\sigma^2}{8(n-2)}} \\ \text{MSE} [\hat{\mu}_3] &= \boxed{\frac{\sigma^2}{n}}\end{aligned}$$

Exercise 9

Let X_1, X_2, \dots, X_n denote a random sample from a distribution with density

$$f(x) = \frac{1}{\theta} e^{-x/\theta}, \quad x > 0, \theta \geq 0$$

Consider five estimators of θ :

$$\hat{\theta}_1 = X_1, \quad \hat{\theta}_2 = \frac{X_1 + X_2}{2}, \quad \hat{\theta}_3 = \frac{X_1 + 2X_2}{3}, \quad \hat{\theta}_4 = \bar{X}, \quad \hat{\theta}_5 = 5$$

Calculate the mean squared error for each estimator. (It will be useful to first calculate their bias and variances.)

Solution:

First, notice that this is an exponential distribution, thus we have

$$\text{E}[X_i] = \theta, \quad \text{Var}[X_i] = \theta^2$$

Following the work from the previous exercise, we will find

$$\begin{aligned} \text{MSE} [\hat{\theta}_1] &= \theta^2 \\ \text{MSE} [\hat{\theta}_2] &= \frac{\theta^2}{2} \\ \text{MSE} [\hat{\theta}_3] &= \frac{5\theta^2}{9} \\ \text{MSE} [\hat{\theta}_4] &= \frac{\theta^2}{n} \\ \text{MSE} [\hat{\theta}_5] &= (5 - \theta)^2 \end{aligned}$$

Note that the first four estimators are unbiased, and thus their mean squared error is simply their variance. The fifth estimator has no variance, and thus its mean squared error is its bias squared.

Exercise 10

Suppose that $E[\hat{\theta}_1] = E[\hat{\theta}_2] = \theta$, $\text{Var}[\hat{\theta}_1] = \sigma_1^2$, $\text{Var}[\hat{\theta}_2] = \sigma_2^2$, and $\text{Cov}[\hat{\theta}_1, \hat{\theta}_2] = \sigma_{12}$. Consider the unbiased estimator

$$\hat{\theta}_3 = a\hat{\theta}_1 + (1 - a)\hat{\theta}_2.$$

If $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent, what value should be chosen for the constant a in order to minimize the variance and thus mean squared error of $\hat{\theta}_3$ as an estimator of θ ?

Solution:

We first verify that $\hat{\theta}_3$ is unbiased.

$$E[\hat{\theta}_3] = E[a\hat{\theta}_1 + (1 - a)\hat{\theta}_2] = aE[\hat{\theta}_1] + (1 - a)E[\hat{\theta}_2] = a\theta + (1 - a)\theta = \theta$$

If $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent, then

$$\text{Var}[\hat{\theta}_3] = \text{Var}[a\hat{\theta}_1 + (1 - a)\hat{\theta}_2] = a^2\text{Var}[\hat{\theta}_1] + (1 - a)^2\text{Var}[\hat{\theta}_2] = a^2\sigma_1^2 + (1 - a)^2\sigma_2^2$$

Thus we want to find a to minimize the function

$$f(a) = a^2\sigma_1^2 + (1 - a)^2\sigma_2^2$$

Taking the derivative, we have

$$f'(a) = 2a\sigma_1^2 - 2(1 - a)\sigma_2^2$$

Setting this equal to 0 and solving for a gives

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

Finally, we check that this gives a minimum by evaluating the second derivative.

$$f''(a) = 2\sigma_1^2 + 2\sigma_2^2$$

Since since this is in fact always positive

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

gives a minimum.

Exercise 11

Let Y have a binomial distribution with parameters n and p . Consider two estimators for p :

$$\hat{p}_1 = \frac{Y}{n}$$

and

$$\hat{p}_2 = \frac{Y + 1}{n + 2}$$

For what values of p does \hat{p}_2 achieve a lower mean square error than \hat{p}_1 ?

Solution:

Recall that for a binomial random variable we have

$$\begin{aligned} \mathbb{E}[Y] &= np \\ \text{Var}[Y] &= np(1-p) \end{aligned}$$

We first calculate the bias, variance, and mean squared error of \hat{p}_1 .

$$\begin{aligned} \text{Bias}[\hat{p}_1] &= \mathbb{E}[\hat{p}_1] - p = p - p = 0 \\ \text{Var}[\hat{p}_1] &= \frac{1}{n^2} \text{Var}[Y] = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n} \\ \text{MSE}[\hat{p}_1] &= \frac{p(1-p)}{n} \end{aligned}$$

Then we first calculate the bias, variance, and mean squared error of \hat{p}_2 .

$$\begin{aligned} \mathbb{E}[\hat{p}_2] &= \mathbb{E}\left[\frac{Y+1}{n+2}\right] = \frac{np+1}{n+2} \\ \text{Bias}[\hat{p}_2] &= \mathbb{E}[\hat{p}_2] - p = \frac{np+1}{n+2} - p \\ \text{Var}[\hat{p}_2] &= \text{Var}\left[\frac{Y+1}{n+2}\right] = \frac{1}{(n+2)^2} \text{Var}[Y] = \frac{np(1-p)}{(n+2)^2} \\ \text{MSE}[\hat{p}_2] &= (\text{Bias}[\hat{p}_2])^2 + \text{Var}[\hat{p}_2] = \left(\frac{np+1}{n+2} - p\right)^2 + \frac{np(1-p)}{(n+2)^2} \end{aligned}$$

Finally, we solve for values of μ where $\text{MSE}[\hat{p}_2] < \text{MSE}[\hat{p}_1]$. We have,

$$\text{MSE}[\hat{p}_2] < \text{MSE}[\hat{p}_1]$$

$$\left(\frac{np+1}{n+2} - p\right)^2 + \frac{np(1-p)}{(n+2)^2} < \frac{p(1-p)}{n}$$

Now we decide to be lazy and use [WolframAlpha](#).

Thus, $\text{MSE}[\hat{p}_2] < \text{MSE}[\hat{p}_1]$ when

$$\boxed{\frac{1}{2} \left(1 - \sqrt{\frac{n+1}{2n+1}}\right) < p < \frac{1}{2} \left(1 + \sqrt{\frac{n+1}{2n+1}}\right)}$$

Note that this interval is symmetric about 0.5, which makes sense because \hat{p}_2 is biased towards 0.5

Exercise 12

Let X_1, X_2, \dots, X_n denote a random sample from a population with mean μ and variance σ^2 .

Create an unbiased estimator for μ^2 . Hint: Start with \bar{X}^2 .

Solution:

First note that from previous results we know that

$$\text{E}[\bar{X}] = \mu, \quad \text{Var}[\bar{X}] = \frac{\sigma^2}{n}$$

Also, recall that re-arranging the variance formula based on the first and second moments, for any random variable we have

$$\text{E}[Y^2] = \text{Var}[Y] + (\text{E}[Y])^2$$

Using these, we find that

$$\text{E}[\bar{X}^2] = \text{Var}[\bar{X}] + (\text{E}[\bar{X}])^2 = \frac{\sigma^2}{n} + \mu^2$$

This expression has the μ^2 that we're after, but also has an extra term that we would like to disappear.

Recall that

$$\text{E}[S^2] = \sigma^2$$

Let's now consider the estimator

$$\hat{\xi} = \bar{X}^2 - \frac{1}{n}S^2$$

It's expected value is

$$E[\hat{\xi}] = E\left[\bar{X}^2 - \frac{1}{n}S^2\right] = E[\bar{X}^2] - \frac{1}{n}E[S^2] = \frac{\sigma^2}{n} + \mu^2 - \frac{1}{n}\sigma^2 = \mu^2$$

Thus $\hat{\xi} = \bar{X}^2 - \frac{1}{n}S^2$ is an unbiased estimator for μ^2 .

Exercise 13

Let $X_1, X_2, X_3, \dots, X_n$ be iid random variables form $U(\theta, \theta + 2)$. (That is, a uniform distribution with a minimum of θ and a maximum of $\theta + 2$.)

Consider the estimator

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

(a) Calculate the **bias** of $\hat{\theta}$ when estimating θ .

Solution:

$$E[\hat{\theta}] = E[\bar{X}] = E[X] = \frac{\theta + (\theta + 2)}{2} = \theta + 1$$

$$\text{bias}[\hat{\theta}] = E[\hat{\theta}] - \theta = (\theta + 1) - \theta = \boxed{1}$$

(b) Calculate the **variance** of $\hat{\theta}$.

Solution:

$$\text{Var}[\hat{\theta}] = \text{Var}[\bar{X}] = \frac{\text{Var}[X]}{n} = \frac{[(\theta + 2) - \theta]^2}{12n} = \boxed{\frac{1}{3n}}$$

(c) Calculate the **mean squared error** of $\hat{\theta}$ when estimating θ .

Solution:

$$\text{MSE}[\hat{\theta}] = (\text{bias}[\hat{\theta}])^2 + \text{var}[\hat{\theta}] = 1^2 + \frac{1}{3n} = \boxed{\frac{3n + 1}{3n}}$$
