

# Homework #8

Winter 2020, STATS 509

Dino Bektesevic

## Problem 1: Properties of Estimators and Confidence Intervals

Goldberger Qu. 11.2

*Hints: For (a) use common sense / the analogy principle; for (b) read Goldberger §10.1, p.107; for part (c) use the analogy principle and p.108; for (d) recall that the standard error of  $T$  is simply an estimate of the standard deviation of  $T$ .<sup>1</sup>*

a. We can pick  $T = \bar{X} - \bar{Y}$  since:

$$E[T] = E[\bar{X}] - E[\bar{Y}] = \mu_X - \mu_Y = \theta$$

b.

$$\begin{aligned} V(T) &= Cov(T, T) = Cov(\bar{X} - \bar{Y}, \bar{X} - \bar{Y}) \\ &= V(\bar{X}) + V(\bar{Y}) - 2Cov(\bar{Y}, \bar{X}) \\ &= V(\bar{X}) + V(\bar{Y}) - 2Cov\left(\frac{1}{n} \sum_i y_i, \frac{1}{n} \sum_i x_i\right) \\ &= V(\bar{X}) + V(\bar{Y}) - \frac{2}{n^2} \sum_{i,j=1}^n Cov(y_i, x_j) \\ &= V(\bar{X}) + V(\bar{Y}) - \frac{2}{n^2} \sum_{i,j=1}^n \sigma_{XY} \\ &= \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{n} - \frac{2}{n^2} n \sigma_{XY} \\ &= \frac{\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}}{n} \end{aligned}$$

c. By analogy with b) we can just say:

$$V(T) = \frac{S_X^2 + S_Y^2 - 2S_{XY}}{n}$$

d. For standard error of  $T$  I would report  $\sqrt{V(T)}$

---

<sup>1</sup>This usage of the term ‘standard error’ follows Goldberger (p.123) who defines the ‘standard error’ of  $\bar{Y}$  to be  $s/\sqrt{n}$  which is an **estimate** of  $\sigma/\sqrt{n}$ , the standard deviation of  $\bar{Y}$ . (Here assuming  $\sigma$  is unknown.)

However, other authors use ‘standard error of  $\bar{Y}$ ’ to refer to  $\sigma/\sqrt{n}$ ; such authors will then refer to  $s/\sqrt{n}$  as an **estimated** standard error. (For Goldberger, adding the word ‘estimated’ to ‘standard error’ would be redundant.)

## Problem 2

Goldberger Qu. 11.3 *Hints: see p.119. For part (a), express the estimator as  $T = a_1\bar{Y}_1 + a_2\bar{Y}_2$ ; find a constraint on  $a_1$  and  $a_2$  in order for  $T$  to be unbiased for  $\mu$ ; use this to solve for  $a_2$  in terms of  $a_1$ ; find the variance of  $T$ ; substitute for  $a_2$ , and then differentiate the variance with respect to  $a_1$ .*

- a. We are asked to consider  $T = c_1\bar{Y}_1 + c_2\bar{Y}_2$ . We want it to be unbiased so:

$$\mu = E[T] = c_1E[\bar{Y}_1] + c_2E[\bar{Y}_2] = c_1\mu + c_2\mu = (c_1 + c_2)\mu$$

we must conclude that  $c_1 + c_2 = 1 \Rightarrow c_2 = 1 - c_1$  if unbiased-ness is to hold true. We are asked to find the minimum variance unbiased estimator so:

$$\begin{aligned} V(T) &= V(c_1\bar{Y}_1 + c_2\bar{Y}_2) \\ &= V(c_1\bar{Y}_1) + V(c_2\bar{Y}_2) \\ &= c_1^2V(\bar{Y}_1) + c_2^2V(\bar{Y}_2) \\ &= c_1^2V(\bar{Y}_1) + (1 - c_1)^2V(\bar{Y}_2) \end{aligned}$$

where we, in line 2, used the fact that the problem tells us that the two samples are independent so  $Cov(\bar{Y}_1, \bar{Y}_2) = 0$ . We minimize the variance as a function of the constants:

$$\begin{aligned} 0 &= \frac{\partial}{\partial c_1} V(T) \\ 0 &= \frac{\partial}{\partial c_1} (c_1^2V(\bar{Y}_1) + (1 - c_1)^2V(\bar{Y}_2)) \\ 0 &= 2c_1V(\bar{Y}_1) - 2(1 - c_1)V(\bar{Y}_2) \\ 0 &= 2c_1(V(\bar{Y}_1) + V(\bar{Y}_2)) - 2V(\bar{Y}_2) \\ c_1 &= \frac{V(\bar{Y}_2)}{V(\bar{Y}_1) + V(\bar{Y}_2)} \\ \rightarrow c_2 = 1 - c_1 &= \frac{V(\bar{Y}_1)}{V(\bar{Y}_1) + V(\bar{Y}_2)} \end{aligned}$$

- b. to verify that  $V(T) < V(\bar{Y}_1), V(\bar{Y}_2)$  we can write:

$$\begin{aligned} V(T) &= c_1\bar{Y}_1 + c_2\bar{Y}_2 \\ &= \left( \frac{V(\bar{Y}_2)}{V(\bar{Y}_1) + V(\bar{Y}_2)} \right)^2 V(\bar{Y}_1) + \left( \frac{V(\bar{Y}_1)}{V(\bar{Y}_1) + V(\bar{Y}_2)} \right)^2 V(\bar{Y}_2) \\ &= \frac{V(\bar{Y}_2)^2V(\bar{Y}_1) + V(\bar{Y}_1)^2V(\bar{Y}_2)}{(V(\bar{Y}_1) + V(\bar{Y}_2))^2} \\ &= \frac{V(\bar{Y}_1)V(\bar{Y}_2)(V(\bar{Y}_2) + V(\bar{Y}_1))}{(V(\bar{Y}_1) + V(\bar{Y}_2))^2} \\ &= \frac{V(\bar{Y}_1)V(\bar{Y}_2)}{V(\bar{Y}_1) + V(\bar{Y}_2)} \end{aligned}$$

To show the inequality we can write

$$\begin{aligned} \frac{V(T)}{V(\bar{Y}_1)} &= \frac{V(\bar{Y}_2)}{V(\bar{Y}_1) + V(\bar{Y}_2)} \leq 1 \\ \frac{V(T)}{V(\bar{Y}_2)} &= \frac{V(\bar{Y}_1)}{V(\bar{Y}_1) + V(\bar{Y}_2)} \leq 1 \end{aligned}$$

since variance is always positive or zero.

### Problem 3

Goldberger Qu. 11.4. Assume that  $Y_1$  and  $Y_2$  are independent.

We are told that we have  $N = 100$  samples where  $n_1$  comes from  $Y_1 \sim N(\mu_1, 50)$  and  $n_2$  comes from  $Y_2 \sim N(\mu_2, 100)$  so that  $N = n_1 + n_2$ .

We are estimating  $T = \mu_1 - \mu_2$  just like in question 1 so we can use  $T = \bar{Y}_1 - \bar{Y}_2$  as an unbiased estimator of  $\theta$ . To get the best estimation of  $\theta$  we want to minimize the variance of  $T$ . So we can write:

$$\begin{aligned} V(T) &= V(\bar{Y}_1) + V(\bar{Y}_2) \\ &= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \\ &= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{N - n_1} \end{aligned}$$

Where we used the fact  $Y_1$  and  $Y_2$  are independent. Minimizing this expression with respect to number of samples  $n_1$  will tell us how many samples of  $Y_1$  we want to get the best estimate of  $\theta$ :

$$\begin{aligned} \frac{\partial}{\partial n_1} V(T) &= \frac{\partial}{\partial n_1} \left[ \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{N - n_1} \right] = 0 \\ -\frac{\sigma_1^2}{n_1^2} + \frac{\sigma_2^2}{(N - n_1)^2} &= 0 \\ \sigma_1^2(N - n_1)^2 &= \sigma_2^2 n_1^2 \\ \sigma_2^2 n_1^2 - \sigma_1^2(N^2 - 2Nn_1 + n_1^2) &= 0 \\ \sigma_2^2 n_1^2 - \sigma_1^2 N^2 + 2\sigma_1^2 Nn_1 - \sigma_1^2 n_1^2 &= 0 \\ (\sigma_2^2 - \sigma_1^2)n_1^2 + 2\sigma_1^2 Nn_1 - \sigma_1^2 N^2 &= 0 \\ \left( \frac{\sigma_2^2}{\sigma_1^2} - 1 \right) n_1^2 + 2Nn_1 - N^2 &= 0 \\ \left( \frac{100}{50} - 1 \right) n_1^2 + 200n_1 - 10000 &= 0 \\ n_1^2 + 200n_1 - 10000 &= 0 \\ \rightarrow n_{1,1} &= 41.42 \\ n_{1,2} &= -241.42 \end{aligned}$$

So we would want to draw 41 sample from  $Y_1$  and 59 from  $Y_2$ .

#### Problem 4: Maximum Likelihood / ZES Estimation / GLRTs

Suppose that  $X_1, \dots, X_n$  are i.i.d. observations from the following pmf:

$$f(x | \theta) = \begin{cases} e^{\theta x} / (1 + e^{\theta}) & x \in \{0, 1\} \\ 0 & \text{otherwise} \end{cases}$$

where  $\theta \in \mathbb{R}$ .

- a. Confirm that for any value of  $\theta$ , this is a probability mass function.

$$\begin{aligned} \sum_x f(x|\theta) &= 1 \\ f(0|\theta) + f(1|\theta) &= 1 \\ \frac{1}{1+e^{\theta}} + \frac{e^{\theta}}{1+e^{\theta}} &= 1 \\ \frac{1+e^{\theta}}{1+e^{\theta}} &= 1 \quad \forall \theta \in \mathbb{R} \end{aligned}$$

- b. Write down the likelihood for one observation  $f(x | \theta)$ . Find the log-likelihood,  $\ell = \log f(x | \theta)$ .
- c. Find the score variable  $Z = (\partial \ell / \partial \theta)$ . Using the fact that  $E[Z] = 0$ , find  $E(X)$  (see Goldberger p.128).
- d. Find the maximum likelihood estimator  $\hat{\theta}$  of  $\theta$  based on the random sample  $X_1, \dots, X_n$ .
- e. Derive the ZES estimator for  $\theta$ . Confirm that this leads to the same estimator for  $\theta$  that you obtained in (d).
- f. Find the asymptotic variance of  $\hat{\theta}$  (this will be a function of  $\theta$ ).
- g. By plugging in  $\hat{\theta}$  for  $\theta$  in your answer to (f), find the standard error of  $\hat{\theta}$ . In other words, find an estimate of the standard deviation of the estimator  $\hat{\theta}$ .
- h. Use your answer to (g) to construct an approximate 95% confidence interval for  $\theta$ . *Hint: make sure that your interval is a function of  $\hat{\theta}$ , NOT the true value of  $\theta$ , which is unknown.*

## Problem 5

Suppose  $Y_1, \dots, Y_n$  are an i.i.d. sample from a population with pmf given by:

$$p(y | \theta) = (y!)^{-1} \theta^y e^{-\theta} \quad (1)$$

where  $\theta > 0$ ,  $y_i \in \{0, 1, \dots\}$ .

- (a) Write down the log-likelihood for a single observation:

$$l(\theta|y) = -\ln(y!) + y \ln(\theta) - \theta$$

- (b) Using your answer to (a) find the score variable for  $\theta$ :

As per Goldbergers definition in chapter 12.1, p 128:

$$Z = \frac{\partial}{\partial \theta} l(\theta|x) = \frac{\partial}{\partial \theta} [-\ln(y!) + y \ln(\theta) - \theta] = \frac{y}{\theta} - 1$$

- (c) Find the information variable for  $\theta$ , and find its expectation:

Information variable as defined in Goldberger 12.2, p 131:

$$W = -\frac{\partial}{\partial \theta} Z = -\frac{\partial}{\partial \theta} \left( \frac{y}{\theta} - 1 \right) = -\frac{y}{\theta^2}$$

Expectation of which:

$$E[W] = \frac{E[y]}{\theta^2}$$

- (d) Find the maximum likelihood estimator  $\hat{\theta}_{MLE}$  for  $\theta$  given the sample  $Y_1, \dots, Y_n$ :

$$0 = \frac{\partial}{\partial \theta} \ln(l(\theta|y_1, \dots, y_n))$$

$$0 = \frac{\partial}{\partial \theta} \ln \left( \prod_{i=1}^n l(\theta|y_i) \right)$$

$$0 = \frac{\partial}{\partial \theta} \sum_{i=1}^n \ln(l(\theta|y_i))$$

$$0 = \sum_{i=1}^n \frac{\partial}{\partial \theta} \ln(l(\theta|y_i))$$

$$0 = \sum_{i=1}^n \frac{y_i}{\theta} - n$$

$$0 = \frac{1}{\theta} \sum_{i=1}^n y_i - n$$

$$n\theta = \sum_{i=1}^n y_i$$

$$\theta_{MLE} = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y}$$

- (e) Using your answers to (c) and (d) give an approximate 90% confidence interval for  $\theta$ : *Hint: your answer should be a function of  $\hat{\theta}_{MLE}$  and  $n$ .*

$$P \left( \bar{Y} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{Y} + 1.96 \frac{\sigma}{\sqrt{n}} \right) = 0.90$$

$$P \left( \theta_{MLE} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \theta_{MLE} + 1.96 \frac{\sigma}{\sqrt{n}} \right) = 0.90$$

## Problem 6

Let  $X_1, \dots, X_n$  be i.i.d. observations from a  $N(\mu, 1)$  population so that  $f(x | \mu) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)^2}$ .

*Hint: See quiz section notes from 12/4/20*

- (a) Find the MLE  $\hat{\mu}_{MLE}$  for  $\mu$ .

Suppose that we wish to perform a likelihood ratio test of the hypothesis  $H_0 : \mu = 0$  against  $H_1 : \mu \neq 0$ .

- (b) Using your answer to (a) write down the generalized likelihood ratio test statistic (LRT).

- (c) Re-express your answer to (b) as a function of  $\bar{X}$ , and draw the LRT as a function of  $\bar{X}$ .

*Hint:  $\sum_{i=1}^n (X_i - \bar{X})^2 = (\sum_{i=1}^n X_i^2) - n(\bar{X})^2$ .*

- (d) If we wish to perform a hypothesis test with significance level  $\alpha = 0.05$ , use your answer to (c) to find the values of  $\bar{X}$  for which we reject  $H_0$ . *Hint: your answer should be a function of  $n$ .*

Suppose that  $n = 100$  and  $\bar{x} = 0.16$ .

- (e) Using your answer to (d), would we reject  $H_0$  in favor of  $H_1$  using significance level  $\alpha = 0.05$ ?

- (f) Find the p-value for this hypothesis test.

## Problem 7

A set of times  $T_1, \dots, T_n$  are sampled independently from a population with the following density:

$$f(t \mid \theta) = \begin{cases} e^{-(t-\theta)} & t \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

where  $\theta > 0$ .

- (a) Find the maximum likelihood estimate for  $\theta$ .

*Hint: do some plots, examining the values of  $\theta$  for which  $f(t_1, \dots, t_n \mid \theta) > 0$ . It may help you first to think about the cases where  $n = 1$  and  $n = 2$ . Do **not** rush into differentiating anything!*

- (b) Is there a ZES estimator for  $\theta$  ? Briefly explain your answer.

[**Motivation:** (not necessary to answer the problem, but may help with intuition). For example, the observations  $T_1, \dots, T_n$  might be the observed times taken for  $n$  messages to be transmitted across a network. In this case,  $\theta$  represents the (non-random) minimum time for a message to be transmitted across the network if there were no delays; the additional random component of the time ( $T - \theta$ ) is due to bottlenecks and queues encountered by the message.]