

# Homework 2 - Stat 495

Justin Papagelis

Due Wednesday, Sept. 21st by midnight

## Practicing Academic Integrity

If you worked with others or used resources outside of provided course material (anything besides our textbook(s), course materials in Moodle, R help menu) to complete this assignment, please acknowledge them below using a bulleted list.

*I acknowledge the following individuals with whom I worked on this assignment:*

Name(s) and corresponding problem(s)

- 

*I used the following sources to help complete this assignment:*

Source(s) and corresponding problem(s)

-

For this assignment, you may do some of the problems on paper if you like, then scan in your solutions and merge with the Integrity page as a cover sheet and any work you are leaving in the .Rmd, such as your work for CASI 5.3. If you decide to type in your answers, you will need to use LaTeX to show your work for some problems.

Because you may merge files into a pdf to submit in any order, be sure to assign pages in Gradescope so I can find work for each problem!

## PROBLEMS TO TURN IN: CASI 2.1, CASI 3.1, Add 1, CASI 4.4 a, Add 2, Add 3, CASI 5.3 (modified)

### CASI 2.1

A coin with probability of heads  $\theta$  is independently flipped  $n$  times, after which  $\theta$  is estimated by

$$\hat{\theta} = \frac{s+1}{n+2};$$

with  $s$  equal to the number of heads observed.  $\hat{\theta}$  will be referred to as the estimator below.

- (a) What are the bias and variance of the estimator?

SOLUTION:

$$E(\hat{\theta}) = E\left(\frac{s+1}{n+2}\right) = \frac{E(s+1)}{E(n+2)} = \frac{E(s)+1}{n+2} = \frac{n\theta+1}{n+2}$$

$$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta = \frac{n\theta+1}{n+2} - \theta = \frac{n\theta+1-n\theta-2\theta}{n+2} = \frac{1-2\theta}{n+2}$$

$$Var(\hat{\theta}) = Var\left(\frac{s+1}{n+2}\right) = \left(\frac{1}{(n+2)^2}\right) Var(s) = \frac{n\theta(1-\theta)}{(n+2)^2}$$

- (b) How would you apply the plug-in principle to get a practical estimate of the standard error of the estimator?

SOLUTION: You could apply the plug-in-principle to get a practical estimate of the standard error by using

$$SE(\theta) = \sqrt{\frac{Var(\theta)}{n}} \text{ and plugging in the observed } \hat{\theta} \text{ for } \theta. \text{ This would yield } SE(\hat{\theta}) = \sqrt{\frac{Var(\hat{\theta})}{n}} = \sqrt{\frac{\frac{n\theta(1-\theta)}{(n+2)^2}}{n}} = \sqrt{\frac{n^2\theta(1-\theta)}{(n+2)^2}}.$$

### CASI 3.1

Suppose the parameter  $\mu$  in the Poisson density (3.3, pg. 23 or 24, depending on your version) is known to have prior density  $g(\mu) = e^{-\mu}$ . What is the posterior density of  $\mu$  given  $x$  (a single observation)?

SOLUTION: The distribution of a single  $x$ :

$$f(x|\mu) = \frac{e^{-\mu} \mu^x}{x!} \propto e^{-\mu} \mu^x$$

The prior density:

$$g(\mu) = e^{-\mu}$$

So the posterior density:

$$g(\mu|x) \propto g(\mu)f(x|\mu) = (e^{-\mu})(e^{-\mu} \mu^x) = e^{-2\mu} \mu^x$$

Which is a Gamma distribution:

$$\mu^x e^{-2\mu} \sim \text{Gamma}(x+1, 2)$$

## Add 1

Suppose you have a random sample of  $n$  observations drawn from a Bernoulli distribution with parameter  $\theta$ . Further suppose that  $\theta$  is unknown, but has a prior density of a Beta( $\alpha$ ,  $\beta$ ) distribution, with both  $\alpha$  and  $\beta$  greater than 0.

part a: Find the posterior density for theta given the data. Be sure to fully specify/identify the posterior density in your solution.

SOLUTION: The distribution of a single  $x$  drawn from Bernoulli distribution:

$$f(x|\theta) = \theta^x(1 - \theta)^{1-x}$$

The joint distribution of all the  $x$ 's:

$$f_n(x|\theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} \propto \theta^x (1 - \theta)^{n-x}$$

The prior distribution:

$$g(\theta) = \left[ \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \right] \theta^{\alpha-1} (1 - \theta)^{\beta-1} \propto \theta^{\alpha-1} (1 - \theta)^{\beta-1}$$

So the posterior distribution is:

$$g(\theta|x) \propto f_n(x|\theta)g(\theta) \propto \theta^x (1 - \theta)^{n-x} \theta^{\alpha-1} (1 - \theta)^{\beta-1} = \theta^{x+\alpha-1} (1 - \theta)^{n-x+\beta-1}$$

Which is a Beta distribution:

$$g(\theta|x) \propto \theta^{x+\alpha-1} (1 - \theta)^{n-x+\beta-1} \sim \text{Beta}(\alpha + x, n - x + \beta)$$

> part b: Find the Bayesian estimator for theta (i.e. the posterior mean).

SOLUTION:

Since  $g(\theta|x) \sim \text{Beta}(\alpha + x, n - x + \beta)$ , the posterior mean for  $\theta$  is

$$\frac{\alpha + x}{\alpha + x + n - x + \beta} = \frac{\alpha + x}{\alpha + n + \beta}$$

### CASI 4.4 a

A coin with unknown probability of heads  $\theta$  is flipped  $n$  times, yielding  $x_1$  heads; then it is flipped another  $x_1$  times, yielding  $x_2$  heads.

(a) What is an intuitively plausible estimate of theta?

SOLUTION: An intuitively plausible estimate of theta is

$$\hat{\theta} = \frac{x_1 + x_2}{n + x_1}$$

.

## Add 2

Suppose  $X_1 \dots X_n$  are iid from an Exponential distribution with parameter  $\beta$ , with pdf:

$$f(x|\beta) = \frac{1}{\beta} \exp^{-x/\beta}, x > 0, \beta > 0,$$

and 0, otherwise.

(a) Find the MLE for beta.

SOLUTION:

$$f_n(x|\beta) = \prod_{i=1}^n \left( \frac{1}{\beta} e^{-x/\beta} \right) = \frac{1}{\beta^n} e^{-(\sum_{i=1}^n x)/\beta}$$

$$\ell(x|\beta) = \log\left(\frac{1}{\beta^n}\right) + \log(e^{-(\sum_{i=1}^n x)/\beta}) = \log(1) - n\log(\beta) - \frac{\sum_{i=1}^n x}{\beta}$$

$$\ell'(x|\beta) = -\frac{n}{\beta} + \frac{\sum_{i=1}^n x}{\beta^2}$$

so

$$-\frac{n}{\hat{\beta}_{mle}} + \frac{\sum_{i=1}^n x}{\hat{\beta}_{mle}^2} = 0$$

$$\frac{n}{\hat{\beta}_{mle}} = \frac{\sum_{i=1}^n x}{\hat{\beta}_{mle}^2}$$

$$\hat{\beta}_{mle} = \frac{\sum_{i=1}^n x}{n} = \bar{x}$$

(b) Verify the MLE is unbiased.

SOLUTION:

$Bias(\hat{\beta}_{mle}) = E(\hat{\beta}_{mle}) - \beta = E(\bar{x}) - \beta = \beta - \beta = 0$ . So the MLE is unbiased.

(c) Find the Fisher information for a single observation.

SOLUTION:

$$f_n(x|\beta) = \frac{1}{\beta} e^{-x/\beta}$$

$$\ell(x|\beta) = \log\left(\frac{1}{\beta}\right) + \log(e^{-x/\beta}) = \log(1) - \log(\beta) - \frac{x}{\beta}$$

$$\ell'(x|\beta) = -\frac{1}{\beta} + \frac{x}{\beta^2}$$

$$\ell''(x|\beta) = \frac{1}{\beta^2} - \frac{2x}{\beta^3}$$

$$E(\ell''(x|\beta)) = E\left(\frac{1}{\beta^2} - \frac{2x}{\beta^3}\right) = \frac{1}{\beta^2} - \frac{2E(x)}{\beta^3} = \frac{1}{\beta^2} - \frac{2\beta}{\beta^3} = \frac{1-2}{\beta^2} = \frac{-1}{\beta^2}$$

$$I_o = -E(\ell''(x|\beta)) = \frac{1}{\beta^2}$$

(d) Find the Cramer Rao lower bound on variance for unbiased estimators of beta.

SOLUTION:

$$I_n(\beta) = nI(\beta) = \frac{n}{\beta^2}$$

so the CRLB is  $\frac{1}{nI_o} = \frac{1}{I_n(\beta)} = \frac{\beta^2}{n}$

$$\text{And } Var(\hat{\beta}_{mle}) = Var(\bar{x}) = Var\left(\frac{\sum_{i=1}^n x}{n}\right) = \frac{1}{n^2} Var(\sum_{i=1}^n x) = \frac{1}{n} Var(x_i) = \frac{\beta^2}{n}.$$

Therefore,  $\hat{\beta}_{mle}$  is the MVUE.

### Add 3

In a few sentences and in your own words, explain what the Neyman-Pearson lemma tells us and why it is important/useful in the context of hypothesis testing.

SOLUTION: The Neyman-Pearson Lemma is a method to determine if the hypothesis test that you are using is the one with the most power at a certain  $\alpha$  level. The Neyman-Pearson Lemma can also be used to construct the most powerful test for  $\alpha$  using the likelihood ratio. This is useful in the context of hypothesis tests because we can create tests that maximize the power of the test against an alternative hypothesis while minimizing the type II error.



### CASI 5.3 (modified)

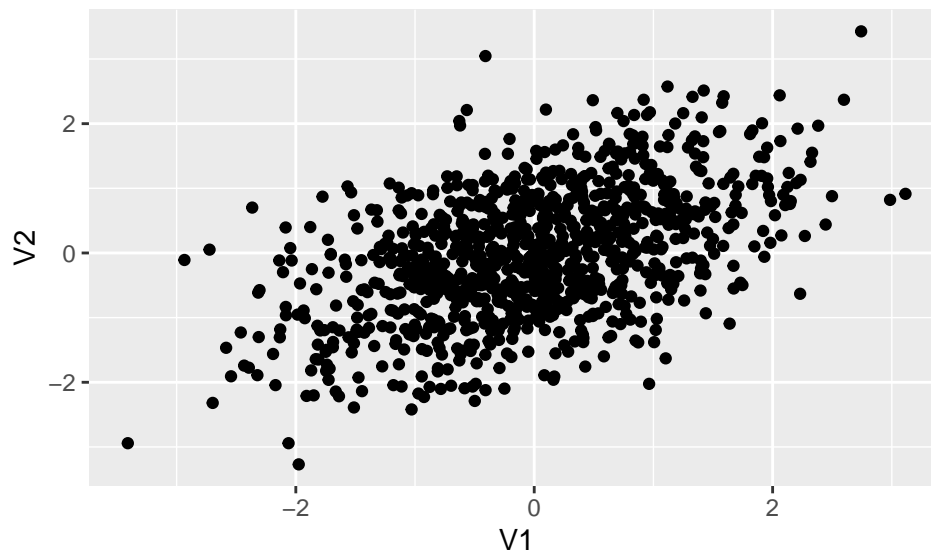
Note: This problem is only visible in the Student edition of the text, not the pdf, but is typed out for you here since I am modifying it anyway.

Draw a sample of 1000 bivariate normal vectors  $x = (x_1, x_2)'$ , with each variable having a mean of 0, a standard deviation of 1, and with a correlation between them of 0.5. Be sure your process is reproducible.

Review the chapter 4 and 5 practice problems for assistance with code.

(a) Plot your sample of data points.

```
set.seed(3)
sigma <- matrix(c(1, 0.5, 0.5, 1), ncol = 2)
NormData <- as.data.frame(rmvnorm(1000, mean = c(0,0), sigma = sigma))
gf_point(V2 ~ V1, data = NormData)
```



```
with(NormData, cor(V1, V2))
```

```
## [1] 0.500141
```

SOLUTION:

(b) Following equation 5.19, what should the theoretical distribution of  $x_2|x_1$  be here?

SOLUTION: The theoretical distribution of  $x_2|x_1$  should be:

$$\begin{aligned}x_2|x_1 &\sim \mathcal{N}\left(\mu_2 + \frac{\sigma_{12}}{\sigma_{11}}(x_1 - \mu_1), \sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}\right) \\x_2|x_1 &\sim \mathcal{N}\left(1 + \frac{0.5}{1}(x_1 - 1), 1 - \frac{0.5^2}{1}\right) \\x_2|x_1 &\sim \mathcal{N}(1 + 0.5x_1 - 0.5, 1 - 0.25) \\x_2|x_1 &\sim \mathcal{N}(0.5x_1 + 0.5, 0.75)\end{aligned}$$

(c) Regress  $x_2$  on  $x_1$  and numerically check equation 5.19.

```
lm1 = lm(V1 ~ V2, data = NormData)
msummary(lm1)
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.000175   0.027511  -0.01    0.99
## V2           0.506563   0.027763  18.25   <2e-16 ***
##
## Residual standard error: 0.87 on 998 degrees of freedom
## Multiple R-squared:  0.25,    Adjusted R-squared:  0.249
## F-statistic: 333 on 1 and 998 DF,  p-value: <2e-16
```

Hint: Reading the text will help you understand what two things to check in the regression output.

SOLUTION: The linear regression coefficient of V2 as a function of V1 is 0.5065 which is close to our theoretical value of  $\frac{\sigma_{12}}{\sigma_{11}\sigma_{22}} = \frac{0.5}{1 \times 1} = 0.5$ . The variance term can also be written as  $\sigma_{22}(1 - R^2)$  which using the  $R^2$  value from the regression output, we get  $1(1 - 0.25) = 0.75$  which is the same variance from our theoretical distribution.