

STAT 517 – Exam 1

Student Gage Sorrell
Professor Dr. Yvonne Zubovich, PhD

Problem #1

It is often based on relevant subject matter and previous studies.

The posterior p.f., $\zeta(\theta|\mathbf{X})$, maps θ as before, but in a way that accounts for newly-observed data, which better informs the statistician regarding the relevant phenomenon. By **Theorem 7.2.1**, it is shown that the posterior p.f. is proportionate to the prior p.f.

The prior p.f. gives the statistician a rough “starting point” to understand the relevant phenomenon, and the posterior p.f. allows the statistician to reach conclusions based on the observed values.

Problem #2

An estimator is a function that produces an estimate, which is an approximation of the true value of the population parameter. The estimator function accepts sample data as input, and can also be a function of unknown parameters.

Problem #3

Part A

The expectation of Karl’s prior is lower than the face-to-face mean score, meaning that Karl believes that online courses are associated with lower scores. Ronald’s expectation is higher, meaning that he believes that online courses are associated with higher scores, *i.e.*, he believes that online courses will be associated with higher scores.

Ronald’s prior has a smaller standard deviation than Karl’s, $\sqrt{8}$ versus $\sqrt{50}$, meaning that he is more confident that scores will be closer to his expectation than Karl is *wrt* his expectation of 65.

Part B

Karl believes that students will perform worse when taking courses online, and Ronald believes that they will perform better. This is modeled in the expectations: $65 < 70$ for Karl, and $80 > 70$ for Ronald. Considering only the mean, one can assume that Karl expects students to only perform slightly worse, whereas Ronald believes that students will perform much better, *i.e.*, by one letter grade.

Taking the variance into account as well shows that Karl expects fewer students to score close to 65 than Ronald expects students to score close to 80, with respective variances 50 and 8.

Problem #4

TODO reword this section.

It is assumed that significant risk of damaging the elevator by overloading it is equally unfavorable as it would be to underload the elevator and forego productivity. Thus, for the loss to be minimized, it should be reflected that the cost of underestimating be roughly equal to the cost of overestimating.

If we choose $c = 1$, then the costs of underestimating and overestimating are equal when $A - \theta = \theta - A \implies A = \theta$. This gives a linear relationship between loss and $A - \theta$.

If we choose $c = 2$, then the cost of underestimating is twice that of overestimating. This is shown by $A - \theta = 2(\theta - A) \implies A = \frac{2}{3}\theta$, and for underestimating, $\theta - A = \theta - \frac{2}{3}\theta = \frac{\theta}{3}$.

If we choose $c = \frac{1}{2}$, then the cost of overestimating is twice that of underestimating. This is shown by $A - \theta = \frac{1}{2}(\theta - A) \implies A = \frac{3}{2}\theta$, and for underestimating, $\theta - A = \theta - \frac{3}{2}\theta = -\frac{\theta}{2}$.

Thus the best choice is $c = 1$.

Problem #5

The statement is not accurate; the MLE of θ does not represent the probability of a particular value of θ . Instead, the MLE of θ maximizes the likelihood function, which represents the probability of observing the given data as a function of θ .

While the concepts in the supposition are closely related, the *causation* implied by the supposition does not reflect how the concepts work.

Problem #6

Recall that the variance of θ is $\text{Var } X = \theta(1 - \theta)$ for a Bernoulli distribution. Using the Method of Moments estimator of $\text{Var } X$, we find the first moment,

$$E(X) = \theta \xrightarrow{\text{MoM}} \theta = \bar{X}_n \implies \hat{\theta} = \bar{X}_n$$

Now we find the second moment, which allows us to solve for the variance,

$$E(X^2) = \text{Var } X + E^2(X) = \text{Var } X + \theta^2 \implies \text{Var } X = E(X^2) - E^2(X) = \theta - \theta^2 = \boxed{\theta(1 - \theta) = \text{Var } X}$$

We plug in the observed data, which allows us to solve for the variance,

$$\hat{\theta} = \bar{X}_n = \frac{160}{200} = 0.8 \implies \text{Var } X = \hat{\theta}(1 - \hat{\theta}) = 0.8(1 - 0.8) = \boxed{0.16}$$

Problem #7

TODO add textbook source here Consider that $E(X) = \int_{-\infty}^{\infty} xf(x|\theta)dx$. We can restrict the interval to where the p.f. is non-zero,

$$E(X) = \int_0^{\theta} \frac{x}{2\theta} dx = \dots = \frac{\theta}{3}$$

Now that we have the population mean, we set the sample mean equal to the population mean,

$$n^{-1} \sum_{i=1}^n X_i = \frac{\hat{\theta}}{3} \implies \hat{\theta} = \frac{3}{n} \sum_{i=1}^n X_i$$

Thus the method of moments estimator is found.

Problem #8

Part A

TODO cite this Recall that by Bayes' theorem,

$$\zeta(\theta|\mathbf{X}) \propto f(\mathbf{X}|\theta)\zeta(\theta) \implies \zeta(\theta|\mathbf{X}) \propto \theta^{-1}\theta e^{-\theta} = e^{-\theta}$$

TODO Cite this Thus by ..., the posterior p.f. is

$$\zeta(\theta|\mathbf{X}) = \begin{cases} \theta e^{-\theta} & \text{for } \theta > 0 \\ 0 & \text{otherwise} \end{cases}$$

Part B

TODO Parts B and C likely have simplified versions in the text.

$$\begin{aligned}
\hat{\theta} &= E(\theta|X) \\
&= \int_0^\infty \theta \cdot \zeta(\theta|X) d\theta \\
&= \int_0^\infty \theta \cdot \frac{\theta e^{-\theta x}}{1 - e^{-\theta x}} d\theta \\
&= \boxed{\int_0^\infty \frac{\theta^2 e^{-\theta x}}{1 - e^{-\theta x}} d\theta}
\end{aligned}$$

Part C

TODO cite this The likelihood function is given by

$$\begin{aligned}
L(\theta|\mathbf{X}) &= \prod_{i=1}^n f(X_i|\theta) \\
&= \prod_{i=1}^n \frac{\theta}{X_i^{\theta+1}} \\
&= \theta^n \left(\prod_{i=1}^n X_i \right)^{-\theta-1}
\end{aligned}$$

Taking the natural logarithm of the likelihood function gives

$$\ln L(\theta|\mathbf{X}) = n \ln \theta - (\theta + 1) \sum_{i=1}^n \ln X_i$$

And the derivative, set to zero

$$\begin{aligned}
\ln_\theta \theta|\mathbf{X} &= \frac{n}{\theta} - \sum_{i=1}^n \ln X_i = 0 \\
0 &= \frac{n}{\theta} - \sum_{i=1}^n \ln X_i \implies \frac{n}{\hat{\theta}} = \sum_{i=1}^n \ln X_i \implies \boxed{\hat{\theta} = \frac{n}{\sum_{i=1}^n \ln X_i}}
\end{aligned}$$

Part B

TODO flesh out solving the integral The first moment is given by

$$\begin{aligned}
E(X) &= \int_1^\infty x f(x|\theta) dx \\
&= \int_1^\infty x \frac{\theta}{x^{\theta+1}} dx \\
&= \theta \int_1^\infty \frac{1}{x^\theta} dx \\
&= \frac{\theta}{\theta - 1}
\end{aligned}$$

We set this equal to the first sample moment,

$$\bar{X}_n = \frac{\theta}{\theta - 1} \implies \bar{X}_n(\theta - 1) = \theta \implies \bar{X}_n\theta - \bar{X}_n = \theta \implies \bar{X}_n\theta = \theta + \bar{X}_n \implies \theta(\bar{X}_n - 1) = \bar{X}_n \implies \hat{\theta} = \frac{\bar{X}_n}{\bar{X}_n - 1}$$