Stat 5102 Notes: Elaborate Solutions to Midterm

Charles J. Geyer

March 3, 2003

General

Generally, notation that *does not appear in the problem statement* should not appear in your solution. If the problem has $x_1, \ldots, x_n$, then you shouldn't have x (and vice versa). If the problem has λ, then you shouldn't have θ. And so forth.

Solve the problem as stated. I know it's much easier to solve problems that are as close as possible to the examples in the book, but how silly is it to say "I can't do any problem in which the data is not denoted by x and the parameter by θ"? You haven't learned how to do statistics if you have a limitation like that.

Problem 1

The main issue here is that if the problem has $x_1, \ldots, x_n$ then so does your solution. The likelihood is

$$\prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta n \bar{x}_n}$$

not

$$\theta e^{-\theta x}$$

If you had the latter, you just weren't being careful enough or didn't know how to be careful about this (weren't aware of the issue).

Problem 2

This problem is much harder than Problem 1. There are dozens of ways to go wrong, and the class exhibited most of them.

First, what's what.

- The data are $x_1, \ldots, x_n$ (not just one x).
- The parameter is λ (not θ or anything else).
- There are no hyperparameters.
- The variables above are the only variables in the problem. No θ, α, β, or anything else should appear in your solution unless you introduce them and define them in terms of λ and $x_1, \ldots, x_n$.

The Likelihood

The data $X_1, \ldots, X_n$ are said to be independent and identically distributed Gamma$(3, \lambda)$.

The Probability Density Function for One X_i

What is the Gamma$(3, \lambda)$ density? If you use my handout on brand name distributions, you see

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha - 1} e^{-\lambda x} \tag{1}$$

is the Gamma(α, λ) density. In this problem $\alpha = 3$, so that gives

$$f(x) = \frac{\lambda^3}{\Gamma(3)} x^2 e^{-\lambda x}$$

The only other thing we need to do here is to write this as a conditional density (so we look like Bayesians)

$$f(x \mid \lambda) = \frac{\lambda^3}{\Gamma(3)} x^2 e^{-\lambda x} \tag{2}$$

The Probability Density Function for One X_i (Alternate)

If you use equation (5.9.7) in DeGroot and Schervish, you see

$$f(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha - 1} e^{-\beta x} \tag{3}$$

In this problem $\alpha = 3$ and $\beta = \lambda$. Doing these substitutions gives (2) again.

The Likelihood

At this point you should have (2) arrived at one way or the other. It should contain x (or x_i) and λ and no other variables.

Since this problem involves n independent and identically distributed variables, the likelihood is the product of n terms like (2)

$$L(\lambda) = \prod_{i=1}^n \frac{\lambda^3}{\Gamma(3)} x_i^2 e^{-\lambda x_i}$$

Any likelihood can be simplified by throwing away *multiplicative* terms that do not contain *the parameter*. In order to use this rule, you must be clear about what multiplicative terms means; the multiplicative terms are those separated by dots below

$$\lambda^3 \cdot \frac{1}{\Gamma(3)} \cdot x_i^2 \cdot e^{-\lambda x_i}$$

And you must be clear about what the parameter is; it is λ. Hence we can throw away $1/\Gamma(3)$ and x_i^2. We cannot throw away λ^3. Many students messed this up.

Hence the likelihood becomes

$$L(\lambda) = \prod_{i=1}^n \lambda^3 e^{-\lambda x_i} = \lambda^{3n} e^{-\lambda \sum_{i=1}^n x_i} \tag{4}$$

Another way to go wrong that several students messed up is bringing the λ^3 out of that product as is instead of as λ^{3n}. Be careful!

The Prior

The parameter λ is said to be distributed Exponential(1). Actually the problem statement said Exp(1), but no one asked or seemed to be confused about what that meant.

If you use my handout on brand name distributions, you see

$$f(x) = \lambda e^{-\lambda x} \tag{5}$$

is the Exponential(λ) density. In this problem the hyperparameter of the prior is 1, so we must make the substitution $\lambda = 1$, so that gives

$$f(x) = e^{-x} \tag{6}$$

Now we need to do two things to look like Bayesians. This is the prior distribution of the *parameter* not the distribution of the *data*. Hence the variable must be λ not x. To distinguish from the data distribution we use another letter g. Thus the prior is

$$g(\lambda) = e^{-\lambda} \tag{7}$$

If you have either (5) or (6) for your prior, then you messed up. A function of x cannot possibly be a function of λ and a prior is a distribution for the parameter λ not data.

In fact there is no x in this problem, only $x_1, \ldots, x_n$. So if you introduced an x here where no x belongs, that by itself should have alerted you to your error.

The Prior (Alternate)

If you use equation (5.9.10) in DeGroot and Schervish, you see

$$f(x \mid \beta) = \beta e^{-\beta x}$$

In this problem $\beta = 1$. Doing this substitution gives (6) again. (We have nothing behind the bar because the hyperparameter β is a known constant.)

The Unnormalized Posterior (Likelihood Times Prior)

Like the heading says, (4) times (7)

$$\lambda^{3n} e^{-\lambda \sum_{i=1}^n x_i} \cdot e^{-\lambda} = \lambda^{3n} e^{-\lambda \sum_{i=1}^n x_i - \lambda} \tag{8}$$

Recognizing the Brand Name Distribution

Admittedly, this is not so obvious, either you see it or you don't. It could be any brand name distribution. Well actually not. The parameter λ is a positive real number. So we are looking at a continuous distribution with sample space $(0, \infty)$. There aren't many of those, the possibilities are

- Exponential
- Gamma
- Chi-Square
- Snedecor's F

The last is out since we haven't learned it yet, it will be a possibility by the second midterm (but isn't a conjugate prior for any brand name data distribution so won't arise in Bayes problems). The exponential and chi-square are special cases of the gamma, so there is really only one choice. The posterior is either gamma or we don't recognize it.

A Possible Posterior

So could (8) be an unnormalized gamma probability density function? To tell we need to be clear about what is the variable. The posterior (like the prior) is a distribution for *the parameter*, in this problem λ. So what does it mean for λ to have a gamma distribution?

If you use my handout on brand name distributions, you see (1) for the Gamma(α, λ) density. Now we want this to be the density of the variable λ, so we will be in big trouble if we have λ as a hyperparameter of the posterior. We would be try to make λ simultaneously represent two different things (parameter and hyperparameter). We must change the λ in (1) to something else, say β, giving (3).

But to repeat, the posterior is a distribution for λ not x, so we must substitute $x = \lambda$ in (3) giving

$$\frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha - 1} e^{-\beta \lambda} \tag{9}$$

If we don't make this substitution, we will be comparing apples and oranges trying to match up (8) and (9).

It helps to see the *form* we are looking for. We can think of (9) as

$$\text{some function of } \lambda = \text{some constant} \cdot \lambda^{\text{something}} e^{-(\text{something else})\lambda}$$

where both "something" and "something else" are positive constants. In (9) "something" is α and "something else" is β, but the letters don't matter, only the form.

So is (8) in this form? Yes! If that isn't obvious, try to make it obvious

$$\lambda^{3n} e^{-\lambda \sum_{i=1}^n x_i - \lambda} = \lambda^{3n} e^{-(\sum_{i=1}^n x_i + 1)\lambda} = \lambda^{\alpha - 1} e^{-\beta \lambda}$$

where

$$\alpha - 1 = 3n$$

so

$$\alpha = 3n + 1$$

(a bit of algebra that I messed up in the first version of the solution) and

$$\beta = 1 + \sum_{i=1}^n x_i$$

To that's the answer: Gamma$(3n + 1, 1 + \sum_{i=1}^n x_i)$.

A Possible Posterior (Alternate)

If you use equation (5.9.7) in DeGroot and Schervish, you see (3) for the Gamma(α, β) density. That puts you midway through the preceding section. As there, you must see the need to substitute $x = \lambda$ giving (9). And from there on the solutions are the same.