Problem Set 2

Jacob M. Montgomery

9/18/2018

The gamma distribution

Let X be a random variable that is distributed according to the gamma distribution such that:

$$X \sim Gamma(\alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$$

where $\Gamma()$ is the gamma function (look it up).

1. Show that this distribution belongs in the exponential family. (Hint: Get *everything* into the exponent.)
2. Assume that α is unknown and β is known. Find the sufficient statistic for α. (Hint: Factorization theorem.)
3. Prove (by direct calculation) that $\frac{\bar{x}}{3}$ is a sufficient statistic for $X \sim Gamma(\alpha = 3, \beta = \theta > 0)$
4. Now let $(X_1, X_2, \ldots, X_n)$ be iid variables drawn from the gamma distribution. Find the Likelihood function.
5. Assume that β is unknown and α is known. Find the Fisher information for β.

Let's try that again

Assume that $(X_1, X_2, \ldots, X_n)$ are drawn from the following pdf

$$f(X|\theta) = \frac{\Gamma(2\theta)}{\Gamma(\theta)^2}[x(1-x)]^{\theta-1}$$

6. Show that this pdf is a member of the exponential family.
7. Find the sufficient statistic for $(X_1, X_2, \ldots, X_n)$.
8. How is your answer to (6) related to your answer to (7). Prove that this relationship will always hold for all members of the exponential family.

Estimating the population variance

9. Let $(X_1, X_2, \ldots, X_n)$ be iid normally distributed data with mean μ and variance σ^2. Show that the sample variance is an unbiased estimator for σ^2.
10. Show that it is a consistent estmator for σ^2.

Bernoulli

11. Prove (by direct calculation) that $\bar{x}$ is a sufficient statistic for $X \sim Bernoulli(\theta)$.
12. Let $X_1, \ldots, X_n \sim Bern(p)$ and let $\hat{p} = \frac{\sum X_i}{n}$. Show that $\hat{p}$ is a consistent estimator for p (ignoring the fact that we have already proved this more generically, but you can follow along using the same basic proof.).

Normal variance

13. Let $X_1, \ldots, X_n$ be iid $N(\mu, \sigma^2)$. The statistic S^2 is an unbiased estimator of σ^2 (you showed this above). Further,
$$E(S^2 - \sigma^2) = Var(S^2) = \frac{2\sigma^4}{n-1}$$
(assertion). Show that S^2 does not attain the information bound.

- Find the log-likelihood.
- Take the second derivative in terms of σ^2.
- Take the expected value and multiply by -1.
- Show that is less than $(\frac{2\sigma^4}{n-1})^{-1}$.

14. Is this estimator efficient (using the finite sample definition)?

And a few more things

15. Find the log likelihood for the normal distribution with unknown mean and variance.
16. Find the second derivative of the log-likelihood in terms of μ and σ^2, noting that the result is going to be a Hessian (and it is going to be a bit ugly). You might want to check your answer with the TA before trying the next part.
17. Remember your defnitions for σ^2 and μ in terms of expectations. Find the Fisher information for the the normal distribution. Lots of things are going to cancel, and the answer is going to look pretty simple (it will be a diagonal matrix). Just do your algebra carefully and do your expectations right.
18. Assume that $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ where neither parameter is known, such that $\theta = (\mu, \sigma^2)$. Use the factorization theorem to show that $\bar{x}$ and $s^2 = \frac{\sum_{i=1}(x_i - \bar{x})}{(n-1)}$ are sufficient statistics for this distribution.
19. Without using the factorization theorem, prove (by direct calculation) that $\bar{x}$ is a sufficient statistic for $X \sim N(\theta, 1)$.
20. Without using the factorization theorem, prove (by direct calculation) that $\frac{\sum x_i^2}{n}$ is a sufficient statistic for $X \sim N(0, \theta)$ (remember: θ is the variance).

---

21. Let $X_1, \ldots, X_n$ by iid logistic where $f(x|\theta) = \frac{e^{-(x-\theta)}}{1+e^{-(x-\theta)^2}}$. Show that the order statistics are a sufficient statistic for θ.