Tufts University
Department of Mathematics
Spring 2022

MA 166: Statistics

Solutions to Homework 4 (v1.0) [1]

Assigned Monday 14 February 2022
Due Monday 21 February 2022 at 11:59 pm EDT.

1. **Larson & Marx, Problem 5.7.3: Suppose $Y_1, Y_2, Y_3, \ldots, Y_n$ is a random sample of size n from the exponential pdf, $f_Y(y;\lambda) = \lambda e^{-\lambda y}$, $y > 0$.**

   (a) **Show that $\hat{\lambda}_n = Y_1$ is not consistent for λ.**
   First note that, for any of the random variables Y_j,
   $$E(Y_j) = \int_0^\infty dy\ f_Y(y;\lambda)y = \int_0^\infty dy\ \lambda e^{-\lambda y} y = \frac{1}{\lambda}\int_0^\infty du\ e^{-u}u = \frac{1}{\lambda},$$
   where we made the u substitution $u = \lambda y$, and integrated by parts or used the straightforwardly proven fact that $\int_0^\infty du\ e^{-u}u^k = k!$. Hence we have
   $$E(\hat{\lambda}_n) = E(Y_1) = \frac{1}{\lambda},$$
   which is not in general equal to λ. Hence the estimator is not unbiased. It is also not asymptotically unbiased, since
   $$\lim_{n\to\infty} E(\hat{\lambda}_n) = E(Y_1) = \frac{1}{\lambda},$$
   which is not in general equal to λ. (That is, the sequence $\{E(\hat{\lambda}_n)\}_{n=1}^\infty$ is a constant sequence, so it converges to its value at $n = 1$.) Since the estimator is not asymptotically unbiased, there is no way that it can be consistent. This is because consistency is a stronger condition than asymptotic unbiasedness. Consistency implies asymptotic unbiasedness, and the contrapositive of that is that lack of asymptotic unbiasedness implies lack of consistency.

   (b) **Show that $\hat{\lambda}_n = \sum_{j=1}^n Y_j$ is not consistent for λ.**
   This time, we have
   $$E(\hat{\lambda}_n) = E\left(\sum_{j=1}^n Y_j\right) = \sum_{j=1}^n E(Y_j) = \sum_{j=1}^n \frac{1}{\lambda} = \frac{n}{\lambda}.$$
   As in part (a), this is not in general equal to λ, hence the estimator is not unbiased. This time it is not asymptotically unbiased because the limit
   $$\lim_{n\to\infty} E(\hat{\lambda}_n) = \lim_{n\to\infty} = \frac{n}{\lambda}$$
   does not even exist. Once again, as explained in the solution to (a), lack of asymptotic unbiasedness implies lack of consistency, so the estimator is not consistent.

2. **In the lecture on Bayesian statistics (Lecture 08), the description of the algorithm for Bayesian searches demonstrates in detail how the prior for region r_j, namely $P(A_j)$, should be updated immediately after an unsuccessful search of region r_j. It does not go into much detail about how $P(A_k)$ should be updated for $k \neq j$ as a result of the same search. Fill in those details to derive the boxed result on slide 13 of that lecture.**

   Referring to the Lecture 08, and using the same notation used in that lecture, after an unsuccessful search in region j, the updated prior for region j is

$$P^*(A_j) = P\left(A_j \mid B_j^C\right) = \left[\frac{1 - P\left(B_j \mid A_j\right)}{1 - P\left(B_j \mid A_j\right)P(A_j)}\right] P(A_j) < P(A_j).$$

   Let's suppose that all the other priors, $P(A_k)$ with $k \neq j$ are scaled by a factor α,

$$P^*(A_k) = \alpha P(A_k) \quad \text{for } k \neq j.$$

   In order for the posterior distribution to be normalized, we must demand that

$$\begin{aligned} 1 =& \sum_{m=1}^{n} P^*(A_m) \\ =& P^*(A_j) + \sum_{k \neq j}^{n} P^*(A_k) \\ =& P^*(A_j) + \alpha \sum_{k \neq j}^{n} P(A_k) \\ =& \left[\frac{1 - P\left(B_j \mid A_j\right)}{1 - P\left(B_j \mid A_j\right)P(A_j)}\right] P(A_j) + \alpha\left(1 - P(A_j)\right), \end{aligned}$$

   and solving for α yields

$$\alpha = \frac{1}{1 - P\left(B_j \mid A_j\right)P(A_j)}.$$

   Hence, the posterior distribution for $k \neq j$ is

$$P^*(A_k) = \left[\frac{1}{1 - P\left(B_j \mid A_j\right)P(A_j)}\right] P(A_k),$$

   as given in the lecture.

3. **Let's suppose that Captain Kidd's long-lost missing treasure has been isolated to three large regions, each of 100 square miles, off of the coast of Rhode Island. The first and third regions are reasonably shallow, so if the treasure is in one of them and it is searched, the probability of finding it is 0.9. The second region is significantly deeper, so if the treasure is there and it is searched, the probability of finding it is only 0.7. Based on historical accounts of Kidd's voyage from the Caribbean to Boston, you would assign**

probabilities of 0.4, 0.1 and 0.5 that the treasure is in each of regions 1, 2 and 3, respectively. What are the first five regions that you would search in a Bayesian search, assuming that all (except possibly the last) turn out to be unsuccessful? Explain in words why searching the regions in this order is sensible. [You may use a computational aid, such as Excel or Mathematica, or a computer language of your choice to do this problem.]

Our probabilities of success if we search in each of the three regions are

$$P\left(\vec{B} \mid \vec{A}\right) = \langle {}^{9}/_{10}, {}^{7}/_{10}, {}^{9}/_{10} \rangle ,$$

where we are using a slightly nonstandard, but obvious, vector notation for the three regions.

At the outset of the problem, our priors are

$$P(\vec{A}) = \langle {}^{4}/_{10}, {}^{1}/_{10}, {}^{5}/_{10} \rangle .$$

The region with the highest probability of success is $\boxed{\textbf{region3}}$, with $P(A_3) = 0.5$, so we search that and we suppose that our search is unsuccessful. The posterior probability for region 3 is then

$$P^*(A_3) = \left[\frac{1 - P\left(B_3 \mid A_3\right)}{1 - P\left(B_3 \mid A_3\right) P(A_3)}\right] P(A_3) = \left[\frac{1 - {}^{9}/_{10}}{1 - {}^{9}/_{10} \times {}^{5}/_{10}}\right] {}^{5}/_{10} = \frac{1}{11} = 0.0909\ldots$$

The factor by which the other priors are scaled is

$$\alpha = \frac{1}{1 - P\left(B_3 \mid A_3\right) P(A_3)} = \frac{1}{1 - {}^{9}/_{10}{}^{5}/_{10}} = \frac{20}{11},$$

so the posterior probabilities for the regions not searched are

$$\begin{aligned} P^*(A_1) =& \alpha P(A_1) = \frac{20}{11} \times \frac{4}{10} = \frac{8}{11} = 0.7272\ldots \\ 410 P^*(A_2) =& \alpha P(A_2) = \frac{20}{11} \times \frac{1}{10} = \frac{2}{11} = 0.1818\ldots \end{aligned}$$

The posterior distribution, after having unsuccessfully searched region 3, is then

$$P^*(\vec{A}) = \langle {}^{8}/_{11}, {}^{2}/_{11}, {}^{1}/_{11} \rangle .$$

The very low posterior probability in region 3 is sensible because there is a high probability of finding the treasure if it is there, and we just searched it and found nothing. The other two regions then scale up accordingly, maintaining their rank order. Because our prior for region 2 was so small, this makes region 1 our primary suspect.

We now treat the posterior distribution calculated above as our prior. From this we see that the next region to search is $\boxed{\textbf{region1}}$, for which our probability is now ${}^{8}/_{11}$, and we again assume that search to be unsuccessful. This time, the scale factor is $\alpha = {}^{55}/_{19}$, and the result for the posterior is

$$P^*(\vec{A}) = \langle {}^{4}/_{19}, {}^{10}/_{19}, {}^{5}/_{19} \rangle .$$

Regions 1 and 3 have now been searched and, in spite of the fact that there is a high probability (0.9) of finding the treasure if it is in both, both searches failed. Hence the probabilities for both of them have now been reduced, but it is worth noting that they are now back to their original ratio of $^4/_5$. This makes region 2, in spite of its low initial prior, the most likely suspect.

Again we treat the posterior distribution calculated above as our prior. From this we see that the next region to search is $\boxed{\textbf{region2}}$, for which our probability is now $^{10}/_{19}$, and we again assume that search to be unsuccessful. This time, the scale factor is $\alpha = {}^{19}/_{12}$, and the result for the posterior is

$$P^*(\vec{A}) = \langle {}^4/_{12}, {}^3/_{12}, {}^5/_{12} \rangle \, .$$

At this point, each region has been searched once, and all three searches failed. We are now back to a state of ignorance even greater than that with which we began. The ratio of the probabilities in regions 1 and 3 is still $^4/_5$, but that for region 2 is no longer much smaller than the other two. It is somewhat smaller, both because the original prior was small, and because there is a greater probability of a false negative in that region. Still, we seem to know less than we thought we did when we started. So region 3 has returned to the status of primary suspect.

Again we treat the posterior distribution calculated above as our prior. From this we see that the next region to search is $\boxed{\textbf{region3}}$, for which our probability is now $^5/_{12}$, and we again assume that search to be unsuccessful. This time, the scale factor is $\alpha = {}^8/_5$, and the result for the posterior is

$$P^*(\vec{A}) = \langle {}^8/_{15}, {}^6/_{15}, {}^1/_{15} \rangle \, .$$

Now that region 3 has been searched twice, its posterior probability is very low compared to the other two, propelling region 1 into the role of primary suspect once again. It is almost the same situation as we had after our very first search, but the probability of region 2 is significantly greater. At this point we can see that there is a kind of periodicity developing – an "echo" of our first searches – so it is not surprising that region 1 will be searched next.

Finally, we treat the posterior distribution calculated above as our prior. From this we see that the next region to search is $\boxed{\textbf{region1}}$, for which our probability is now $^8/_{15}$, and we again assume that search to be unsuccessful. This time, the scale factor is $\alpha = {}^{25}/_{13}$, and the result for the posterior is

$$P^*(\vec{A}) = \langle {}^4/_{39}, {}^{30}/_{39}, {}^5/_{39} \rangle \, .$$

After two unsuccessful searches each of regions 1 and 3, region 2 has emerged as the frontrunner once again.

Our search sequence for the first five searches has been 3, 1, 2, 3, 1. If we were to keep this up, here is how this sequence would continue (including the first five):

$$\begin{aligned}\{&3,1,2,\\ &3,1,2,2,3,1,2,2,3,1,2,2,\\ &3,1,2,2,3,1,2,2,3,1,2,2,\\ &3,1,2,3,2,1,2,3,2,1,2,3,\\ &1,2,2,3,1,2,2,3,1,2,2,\ldots\}\end{aligned}$$

It is evident that this sequence is *not* periodic, though it does fall into patterns. After two rounds of searching in the $\{3,1,2\}$ pattern, it falls into a $\{3,1,2,2\}$ pattern for a while, presumably to compensate for the lower certainty of finding the treasure in region 2. Then, after a somewhat chaotic sequence of searches in the fourth row above, it falls into a $\{1,2,2,3\}$ pattern. The probability distributions never seem to recur – at least not for as long as I have looked at it – so it is never exactly periodic. The behavior is fascinating, and seems to be worthy of further study.

4. **Larson & Marx, Problem 5.8.1: Suppose that X is a geometric random variable, where $p_X(k|\theta) = (1-\theta)^{k-1}\theta$, $k = 1,2,\ldots$. Assume that the prior distribution for θ is the beta pdf with parameters r and s. Find the posterior distribution for θ.**

The beta pdf is defined on page 332 as

$$f_\Theta(\theta) = \frac{1}{B(r,s)}\theta^{r-1}(1-\theta)^{s-1},$$

where the beta function is defined to be

$$B(x,y) := \int_0^1 d\xi\ \xi^{x-1}(1-\xi)^{y-1} = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

so the pdf is normalized. Note that the beta function is symmetric in its two arguments, $B(x,y) = B(y,x)$. This follows from the integral by the substitution $u = 1-\xi$, and it is manifest from the expression in terms of the gamma functions. The relationship between the beta and gamma functions is not at all obvious, but the proof is not difficult and is presented nicely in the Wikipedia page on the beta function.

Because X is a discrete random variable and $\Theta \in [0,1]$ is a continuous random variable, the expression for g_θ given at the top of page 332 in Larsen & Marx becomes

$$\begin{aligned} g_\Theta(\theta \mid X=k) =& \frac{p_X(k|\theta)f_\Theta(\theta)}{\int_0^1 d\xi\ p_X(k|\xi)f_\Theta(\xi)} \\ =& \frac{(1-\theta)^{k-1}\theta\left[\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\theta^{r-1}(1-\theta)^{s-1}\right]}{\int_0^1 d\xi\ (1-\xi)^{k-1}\xi\left[\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\xi^{r-1}(1-\xi)^{s-1}\right]} \\ =& \frac{\theta^r(1-\theta)^{k+s-2}}{\int_0^1 d\xi\ \xi^{(r+1)-1}(1-\xi)^{(k+s-1)-1}} \\ =& \frac{\theta^r(1-\theta)^{k+s-2}}{B(r+1,k+s-1)}, \end{aligned}$$

whence

$$\boxed{g_\Theta\left(\theta \mid X=k\right) = \frac{\theta^r(1-\theta)^{k+s-2}}{B(r+1,k+s-1)} = \frac{\Gamma(r+s+k)}{\Gamma(r+1)\Gamma(k+s-1)}\theta^r(1-\theta)^{k+s-2}}$$

This result may be compared with the prior

$$f_\Theta(\theta) = \frac{\theta^{r-1}(1-\theta)^{s-1}}{B(r,s)} = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}\theta^{r-1}(1-\theta)^{s-1}.$$

It is seen that the parameters in the prior beta distribution have changed due to the observation that $X=k$, but the posterior distribution is still a beta distribution.