

STA 243 Homework 4

Ariel Sim (50% of work) Samantha Maillie (50% of work)

5/14/2018

Problem 1

Method 1:

In these three problems there are actually two methods that we compared. We know the true value values are .33333 for a, 3.48318 for b, and 2.27477 for c. A very simple preliminary method we can use is the “dart board method”. The goal is to create a rectangle or box that encloses the entire function. The plots are displayed so that the axes represent the lines of the rectangle/box. This method works very well for a and c. They both converge well within 999 iterations to the true value. The 2 dimensional process is randomly generate (x,y) coordinates in the rectangle. Then the proportion that fall under the curve is multiplied by the area of the rectangle and this gives us our estimate. This method does not perform as well for part b. Now we generate (x,y,z) coordinates and find the proportion that fall below the function plane. I have previously used this method and had decent results. I believe this particular function illustrates how this method can quickly run into issues. It takes about 100,000 iterations for the estimation to get close to 3.4 and even at that it is underestimating the volume under the plane to be about 3.2 as opposed to 3.4. My guess is a few things are at play that are causing issues in this rough method. One is the function is not by any means a simple shape. Just figuring out the dimensions of the enclosing box were difficult. We have to be careful to remain within the bounds of x and y. There could be some rounding error on where the box needs to end. Then we also have a cross over the z-axis which complicates the problem further. The conclusion is that while this works well for a nice and clean function the second method we looked at is more dependable and therefore better for use in real life applications since we wouldn't know the true error.

The 2nd method that I used for (b) instead assumes that $X \sim \text{Unif}[-2,2]$ and $Y \sim \text{Unif}[0,1]$ and are mutually independent. Then, we can factorize $h^*(x,y) = f(x,y)h(x,y) = x^2 \cos(xy)$, where $f(x,y) = f(x)f(y) = 1/4$ and therefore, $h(x,y) = 4x^2 \cos(xy)$. So, by generating n X and Y values using `runif()` in R and then estimating the integral using the mean of the h function, we get an estimate that is ~ 3.53 , which is much reasonably closer to 3.48318

Problem 2

The clear selection of v is .1 Both .1 and 1 do decent jobs of calculating the true integral value which is approximately .1359. .1 has a much smaller standard error though. $v = 10$ is consistently doing a very poor job. It has a small error but a small error for a consistently bad estimation.

Problem 3

- (a) The estimate for $I_{MC}^{\hat{}}$ is 0.6934366 which is relatively close to $\ln(2) = 0.6931472$.
- (b) $E[c(U)] = E([1 + U]) = 1 + 0.5$. The optimal $b \approx 0.001631918$ using the formula provided:

$$b = \frac{\text{Cov}(I_{MC}^{\hat{}}, \theta_{MC}^{\hat{}})}{\text{Var}(\theta_{MC}^{\hat{}})}$$

and calculations for numerator and denominator can be found in the code provided.

- (c) Since we've computed the covariances and variances in (b), we can find $\text{Var}(\hat{I}_{CV}) = \text{Var}(\hat{I}_{MC}) + b^2 \text{Var}(\hat{\theta}_{MC}) - 2b \text{Cov}(I_{MC}^{\hat{}}, \theta_{MC}^{\hat{}}) = 4.16\text{e-}07$. The sample variance of $\hat{I}_{MC} = 1.31\text{e-}05$. The variance of the estimator is smaller after applying the control variate.
- (d) To get a smaller variance, we can change the mapping of $c = 1 + x$ to $c = 1 + x^{1/2}$.

Problem 4

- (a) The model is then $y_{ij} = \mu_i + \epsilon_{ij}$, $i = 1, 2, 3$ for each treatment and the residuals have iid double exponential distributions centered on zero as stated in the problem. The Monte Carlo method for testing for the main effects of each treatment is as follows:
1. Generate Y_{i1}, \dots, Y_{in} from the model as stated in the problem for each treatment, $i = 1, 2, 3$.
 2. Calculate the treatment means, $\bar{Y}_i = \frac{\sum Y_{ij}}{n_i}$ for each i 's.
 3. Repeat Steps 1 and 2 for 999 times.
 4. For each i , reject $H_0 : \mu_i$'s are all equal i.e. there is no treatment effect if the calculated \bar{Y}_i is amongst the smallest or largest 2.5% of the generated Y_{ij} 's for each i .
- (b) If we make no assumptions about the residuals, then we will use Fisher's permutation test. As such, the steps are as follows:
1. Define the test statistic $\hat{\theta} = \bar{Y}_i - \bar{Y}_{i'}$ for $i = 1, 2, 3$ and $i \neq i'$. The null hypothesis is then $\hat{\theta} = 0$ and alternative $\hat{\theta} \neq 0$.
 2. For each $i \neq i'$, combine the observed values for i and i' to get $n_i + n_{i'}$ total observations for sampling in Step 3.
 3. Sample n_i points without replacement as group 1, and then remainder is group 2.
 4. Calculate the sample mean for each group and take their difference, $\hat{\theta}^*$.
 5. Repeat steps 2-4 999 times. If $\hat{\theta}$ falls outside the 95% mean differences of the 999 calculated $\hat{\theta}^*$'s, then we reject the null hypothesis. Otherwise, do not reject.
 6. If we can reject at least one of these differences in means, then we conclude there is a treatment effect. Otherwise, we conclude that no treatment effect is present.

Problem 5

- (a) See code
- (b) These three densities are proportional to $f(x, r, \lambda, p)$. An important thing to note is since we are working with proportionality and not equivalence we can ignore terms not related to the specific variable we are looking at since they will behave like constants.

i. $(\lambda|r, r, x) \propto \lambda^{a-1} e^{-b\lambda} \prod e^{-\lambda r_i} (\lambda r_i)^{x_i}$

If we expand on the product and rearrange the terms with the goal distribution in mind we see the following:

$$(\lambda|r, r, x) \propto \lambda^{a-1+\sum x_i} e^{-[b+\sum r_i]\lambda}$$

This is proportional to the gamma distribution $\text{Gamma}(a-1+\sum x_i, b+\sum r_i)$

$$(\lambda|r, r, x) \propto \frac{(b+\sum r_i)^{a+\sum x_i} \lambda^{a-1+\sum x_i} e^{-[b+\sum r_i]\lambda}}{\Gamma(a+\sum x_i)}$$

- ii. Again we can ignore any terms not related to p and then we can rearrange with the goal distribution in mind.

$$(p|\lambda, r, x) \propto \prod p^{r_i} (1-p)^{1-r_i}$$

$$(p|\lambda, r, x) \propto p^{\sum r_i} (1-p)^{n-\sum r_i}$$

This is clearly proportional to the beta distribution.

$$(p|\lambda, r, x) \propto \frac{p^{\sum r_i} (1-p)^{n-\sum r_i}}{B(1+\sum r_i, n+1-\sum r_i)}$$

- iii. We repeat the same process.

$$(r_i|\lambda, p, x) \propto \sum e^{-\lambda r_i} (\lambda r_i)^{x_i} p^{r_i} (1-p)^{1-r_i}$$

$$(r_i|\lambda, p, x) \propto e^{-\lambda \sum r_i} p^{\sum r_i} (1-p)^{n-\sum r_i} \prod (r_i)^{x_i}$$

Remember that all but r acts as a constant here.

$$(r_i|\lambda, p, x) \propto \left(\frac{pe^{-\lambda}}{pe^{-\lambda} + (1-p)I_{r_i=0}} \right)^{\sum r_i} \left(1 - \frac{pe^{-\lambda}}{pe^{-\lambda} + (1-p)I_{r_i=0}} \right)^{n-\sum r_i}$$

- (c) The 95% confidence intervals can be constructed using the 2.5 and 97.5 quantiles. It is clear from the plotted histogram that the method is estimating p and λ very well. It is not a surprise that both confidence intervals contain what we know to be the true value. The true value of p is .3 with a calculated 95% confidence interval of (.2404, .4921). The true value of λ is 2 with a calculated 95% confidence interval of (1.047, 2.116).

Problem 6

This is a pretty straight forward method. We tried some different shapes and rates but ultimately decided to stick with shape = 2 and rate = 1. It tends to estimate the $E[1/x]$ almost exactly right and underestimates a little but not much for the $E[x]$.