EXERCISE ON BAYESIAN BASICS

We have a random sample of N people who all became unemployed at the same date $t = 0$. For each person i the number of months that he or she remained unemployed before taking a new job is X_i. We believe X_i is independent across i and identically distributed, with pdf $\alpha e^{-\alpha x_i}$, i.e. exponentially distributed. α is unknown.

We know that the people in the sample were observed over just 10 months, and that for anyone who did not have a job at that point the observed data Z show $Z_i = 10$. For everyone who does find a job within 10 months, $Z_i = X_i$.

(1) Display the likelihood function (which will depend on the observed sample of Z values and α, but not on X values).

$$\alpha^n \exp\left(-\alpha \sum Z_j\right) ,$$

where n is the number of observations with $Z_j < 10$.

(2) Show that the sample mean of X_i, if we could see it, would be an unbiased estimate of $1/\alpha$, but that the sample mean of Z_i is not unbiased for $1/\alpha$.

If we could see X_i, the sample pdf would be

$$\alpha^N \exp(-\alpha \sum X_j) .$$

Here you could just invoke the fairly well known fact that a sum of N independent $\mathrm{Gamma}(m, \alpha)$ variates is $\mathrm{name}(mN, \alpha)$. But one could also directly transform from $\{X_1, \dots, X_N\}$ to $\{Y, X_2, \dots, X_N\}$, where $Y = \sum(X_j)$, and integrate out $\{X_2, \dots, X_N\}$. The Jacobian of the transformation is 1, and the integration must take into account that, e.g., $X_N \in (0, Y - \sum_1^{N-1} X_j)$. The expected value of a $\mathrm{Gamma}(mn, \alpha)$ variate is mn/α. In our case $m = 1$, which leads to the unbiasedness result.

In any sample the sample mean of $\{Z_i\}$ is less than or equal to the sample mean of $\{X_i\}$, and it is strictly less when any $Z_i = 10$ is in the sample. Since the sample will have some $Z_i = 10$ occurrences with positive probability, the expected value of the sample mean of the Z_i's is strictly less than the expected value of X_i.

(3) For each of the two sets of sample values below

(a) Calculate the flat-prior Bayesian posterior mean for α and a 90% HPD interval for it. Plot the posterior pdf.

The likelihood as a function of α is proportional to a $\mathrm{Gamma}(n + 1, \sum Z_j)$ pdf, so the posterior mean is $(n + 1)/ \sum Z_j$, where, as above, n is the number of observations with $Z_j < 10$. So the posterior means for the two samples are .149 and .314, respectively. The hpd intervals for the two samples are (.0646, .2305) and (.154, .470), respectively. (Plots omitted since we saw those in class)

(b) Note that in the second sample every person has found a job in less than 10 months. The Bayesian posterior mean is therefore the same as it would have

Date: January 15, 2018.

been if people had been observed longer. Is Bayesian inference, including the HPD interval, therefore with that sample completely unaffected by the presence of the 10-month truncation?
Yes.

(c) For the second sample, since the sample average of the Z_i's also the sample average of the X_i's, can we say that it is unbiased for $1/\alpha$ for this case?
No. Unbiasedness is a property that holds as an average across potential samples. The fact that this sample has no 10's doesn't change the fact that 10's are possible in other samples, and this possibility means the sample mean of $Z's$ is downward biased. Of course for decision-making purposes, the posterior mean under a (nearly) flat prior will be (nearly) the mean of the Z's for this sample, which may be closer to what most people would think "unbiasedness" means than the technical frequentist definition of that term.

(d) Compute a 90%confidence interval for α for each of the two samples. This probably has to be computationally intensive. One approach would be to construct a 90% test of $H0 : \alpha = a$ for a grid of a values and then collect all values not rejected as your interval. Another approach would be to treat the mean of $\sum z_i$ as approximately normal, with mean and variance a known function of α. This would be only asymptotically justfied, but we might see how far off it is for these samples and this sample size.
Code to do this based on likelihood ratio tests is below. The intervals are (.072, .256) and (.17, .518) — quite close to the HPD intervals.

The data below are also available as a .txt file or a .RData file on the web site.

sample 1	sample 2
3.82	0.43
9.23	4.57
1.91	0.15
0.16	1.48
10.00	3.52
1.66	6.27
1.18	6.11
7.05	7.37
10.00	0.49
8.66	1.41

Code for HPD interval:

```
hpd <- function(prob, n, s) {
    uplim <- qgamma(1-prob-.001, n, s)
    ddif <- function(a) {
        b <- qgamma(pgamma(a, n, s) + prob, n, s)
        dgamma(a, n, s) - dgamma(b, n, s)
    }
    x <- uniroot(ddif, c(0, uplim))
    ## the uniroot function in R is just a one-d zero-finder
    y <- qgamma(pgamma(x$root, n, s) + prob, n, s)
    return(c(x,y=y))
}
```

Code for LR confidence interval:

```
n <- matrix(0,1000,1000, dimnames=list(draw=NULL, alpha=NULL))
sz <- matrix(0,1000,1000, dimnames=list(draw=NULL, alpha=NULL))
alpha <- 1:1000/500
X <- matrix(rgamma(10000, 1), 1000, 10, dimnames=list(draw=NULL, j=NULL))
rjct <- vector("numeric", 1000)
for(aj in 1:1000) {
    Z <- X/alpha[aj]
    n <- apply(Z, 1, function(w) sum(w < 10))
    sz <- apply(Z, 1, function(w) sum(pmin(w,10)))
    lr <- ifelse(n > 0, n * log(alpha[aj] / n) + n * log(sz) + n, 0)
     - alpha[aj] * sz
    rjct[aj] <- sort(lr)[100]
}
sz1 <- sum(z1)
llr1 <- 8 * log(alpha/8) + 8 * log(sz1) + 8 - alpha * sz1
min(alpha[llr1 > rjct])
max(alpha[llr1 > rjct])
sz2 <- sum(z2)
llr2 <- 10 * log(alpha/10) + 10 * log(sz2) + 10 - alpha * sz2
min(alpha[llr2 > rjct])
max(alpha[llr2 > rjct])
```