

# Assignment 4

Harry Bendekgey  
Math 153: Bayesian Statistics

March 29, 2018

1. Recall the definition of the frequentist p-value: the probability, under the null hypothesis, of seeing data as or more contradictory to the null hypothesis as our data.

Consider the following hypothesis test:

$$H_0 : p \geq .5 \quad H_1 : p < .5$$

Suppose that some years later, while cleaning out the lab of a retired scientist, we find the results of a particular experiment. A total of 12 trials were conducted, and only 3 successes were observed. Assume that the experiment was conducted by deciding beforehand (based maybe on time or financial constraints) that 12 trials would be collected.

- a. Write down the set of values of  $x$  that constitute “as or more contradictory to the null” (assuming the test given above).

Because we are expecting  $p$  to be large, the smaller  $x$  is the more contradictory it is to the null. Thus, any experimental results where  $x \leq 3$  are as or more contradictory to the null than what was observed.

- b. Compute the p-value for the test. (Use R.) Under the standard significance level of  $\alpha = .05$ , are you publishing a paper with your retired colleague?

```
pbinom(3, size=12, prob=.5)
```

```
## [1] 0.07299805
```

We get a p-value of .07, which is not small enough to reject the null under the standard significance level.

- c. Suppose that in conversation with another senior colleague, you mention the experiment, and she recalls that the data was actually collected by deciding that trials would be conducted until three successes had been observed. Compute the p-value now (again, use R). Are you publishing a paper now?

```
1-pnbinom(9, size=3, prob=.5)
## [1] 0.01928711
```

In this case, as or more contradictory means more than 9 failures needed before 3 successes are seen. In this case, we get a p-value of 0.02, which is small enough to reject the null. We're published!

d. How does the above make you feel?

This hurts me deeply.  
My heart aches for better math.  
Stupid frequentists.  
-a haiku, probably

e. Assuming that a prior is available, explain why the results would be the same in (b) and (c) in a Bayesian setting.

When using Bayes' Rule, we can throw out all constants and look at the kernel of the product of the likelihood and the prior. These constants will cancel with the marginal so we don't care about them. Both the binomial and negative binomial have the same kernel with respect to  $p$ :  $p^k(1-p)^{n-k}$ . The difference is in the constants, which again will get canceled out with the marginal. Thus, under both scenarios the posterior will be exactly the same, which is what we'd want.

2. Let  $y|\theta \sim \text{Pois}(\theta)$  and assume a prior on  $\theta$  of  $\text{Gamma}(1,1)$ .

a. For  $H_0 : \theta \leq 1$  versus  $H_1 : \theta > 1$ , obtain the formula for the posterior probability that  $H_0$  is true. Calculate this for 3 different samples of size 1:  $y = 3, 5, 7$ .

Using Bayes' rule, we see:

$$f(\theta|y) \propto \theta^y e^{-\theta} \cdot \theta^0 e^{-\theta} = \theta^y e^{-2\theta}$$

We recognize this as the kernel of a gamma distribution with parameters  $\alpha = y + 1$  and  $\beta = 2$ . The probability of the null hypothesis is simply the integral from 0 to 1 of the posterior, which gives us:

$$\int_0^1 \frac{2^{y+1}}{\Gamma(y+1)} \theta^y e^{-2\theta} d\theta$$

```
pgamma(1, shape=4, rate=2)
## [1] 0.1428765
pgamma(1, shape=6, rate=2)
```

```
## [1] 0.01656361

pgamma(1,shape=8,rate=2)

## [1] 0.001096719
```

b. For  $H_0 : \theta = 1$  versus  $H_1 : \theta \neq 1$ , with  $q_0 = P(H_0) = 0.5$ , and  $p_1(\theta) = e^{-\theta}$ , compute the posterior probability that  $H_0$  is true for the three  $y$  values. Using Bayes' Rule, we see:

$$P(H_0|y) = \frac{P(y|H_0)P(H_0)}{P(y|H_0)P(H_0) + P(y|H_1)P(H_1)}$$

We note that  $P(H_0) = P(H_1) = 0.5$ , and that  $P(y|H_0)$  is just the likelihood of the Poisson distribution.

To calculate  $P(y|H_1)$  we use the law of total probability:

$$P(y|H_1) = \int_0^\infty p(y|\theta)p(\theta)d\theta = \int_0^\infty \frac{\theta^y e^{-\theta}}{y!} e^{-\theta} = \frac{1}{y!} \int_0^\infty \theta^y e^{-2\theta} d\theta$$

We recognize this to be the kernel of a gamma distribution with  $\alpha = y + 1$  and  $\beta = 2$ . It therefore integrates to the inverse of the normalizing constants,  $\frac{y!}{2^{y+1}}$ . The factorials cancel and we get  $P(y|H_1) = \frac{1}{2^{y+1}}$ . The  $P(H_0)$  and  $P(H_1)$  which are all 0.5 cancel. Thus we get:

$$P(H_0|y) = \frac{\text{Pois}_1(y)}{\text{Pois}_1(y) + 0.5^{y+1}}$$

```
dpois(3,lambda=1)/ (dpois(3,lambda=1)+ 0.5^(4))

## [1] 0.4952075

dpois(5,lambda=1)/ (dpois(5,lambda=1)+ 0.5^(4))

## [1] 0.04675713

dpois(7,lambda=1)/ (dpois(7,lambda=1)+ 0.5^(4))

## [1] 0.001166509
```

3. Frequentist hypothesis tests have poor performance when the sample size is very large (not uncommon) in that they will tend to always reject the null hypothesis, even when the difference between the null hypothesis and the truth is not practically significant.

a. Suppose we are testing for the fairness of a coin, that we think may be biased towards heads but is certainly not biased towards tails. Thus, we are thinking

of the test  $H_0 : p = .5$  vs  $H_1 : p > .5$ . Via the central limit theorem, you may recall that we will reject the null if  $\hat{p} > .5 + 1.645\sqrt{\frac{.25}{n}}$ . Find the sample size that will reject with 95 percent probability (power) a coin with  $p = .501$ . We want to find the sample size at which a binomial distribution centered at .501 has 95% of its area greater than the probability cutoff of a corresponding distribution centered at 0.5. We use the normal approximation (and approximate the standard deviation of the distribution centered at 0.501 to be  $\sqrt{\frac{.25}{n}}$ :

$$\begin{aligned} .5 + 1.645\sqrt{\frac{.25}{n}} &= .501 - 1.645\sqrt{\frac{.25}{n}} \\ \sqrt{\frac{.25}{n}} &= \frac{.001}{3.29} \\ n &= \frac{3290^2}{4} = 2706025 \end{aligned}$$

b. Suppose that we would only actually be interested in the result if  $p \geq .55$ . We might consider using a normal prior centered at .5 as the distribution of  $p|H_0$ , and a *Beta* distribution that gives very little probability less than .55 for the distribution of  $p|H_1$ . Specify these priors, and assuming that we saw  $X = 527$  in  $n = 1000$ , use stochastic integration to compute the Bayes Factor.

$$BF_{10} = \frac{P(y|H_1)}{P(y|H_0)}$$

So sample from the null and alternative distributions repeatedly, computing the likelihood for each value of  $p$ , and then average these to get an estimate of the integrals.

Law of total probability tells us:

$$\frac{P(y|H_1)}{P(y|H_0)} = \frac{\int f(y|\theta, H_1)f(\theta|H_1)}{\int f(y|\theta, H_0)f(\theta|H_0)} = \frac{\int \binom{1000}{527} p^{527} (1-p)^{473} f(\theta|H_1)}{\int \binom{1000}{527} p^{527} (1-p)^{473} f(\theta|H_0)}$$

Now we recognize the numerator to be  $E[g(p)]$  with regard to the alternative prior, where  $g(p)$  is just the pmf of a binomial distribution with  $n = 1000$  and  $p = p$  evaluated at 527. The denominator is the same with regard to the null prior. Thus we can repeatedly sample from the alternative prior, plug it into that equation, and find the average of those values, and the law of large numbers will give us the numerator. We can do the same for the denominator. In fact we don't need to average them, we can just sum them assuming we make the same number of draws in both, because the division by  $n$  cancels.

Finally, we need to specify our priors. For our beta density, let's choose Beta(10,2). It has about 1% of its distribution below 0.55. For our normal, let's use Normal(.5,.15).

```

num.gens <- rbeta(1000,10,2)
denom.gens <- rnorm(1000,.5,.1)
nums <- c()
denoms <- c()
for (i in 1:1000) {
  nums[i] <- dbinom(527, size=1000, prob=num.gens[i])
  denoms[i] <- dbinom(527, size=1000, prob=denom.gens[i])
}
mean(nums)/mean(denoms)

## [1] 0.03666178

```

There is absolutely no evidence to contradict the null. This value has a negative log, putting it even below “not worth more than a bare mention”.  
c. Compute the frequentist p-value for the above data (you can use *prop.test*).

```

prop.test(x=527,n=1000,alternative="greater")

##
## 1-sample proportions test with continuity correction
##
## data: 527 out of 1000, null probability 0.5
## X-squared = 2.809, df = 1, p-value = 0.04687
## alternative hypothesis: true p is greater than 0.5
## 95 percent confidence interval:
## 0.5004926 1.0000000
## sample estimates:
## p
## 0.527

```

We reject the null hypothesis! We're published!