STAT 323 Assignment 5

1. 1

   We set our two tests equal:

$$\begin{aligned} P(Y_1 > .95|\theta = 0) &= P(Y_1 + Y_2 > c) \\ \int_{0.95}^{1} 1\,dx &= P(Y_1 + Y_2 > c) \\ 0.05 &= P(Y_1 + Y_2 > c). \end{aligned}$$

   Now from example 1.1.6, $Y_1 + Y_2 = Z$ has the pdf given by:

$$f_Z(z) = \begin{cases} z & 0 \le z \le 1 \\ -z + 2 & 1 < z \le 2 \\ 0 & \text{otherwise} \end{cases}.$$

   And so we integrate;

$$\begin{aligned} 0.95 &= P(Z \le c) \\ &= \int_0^1 z\,dz + \int_1^c -z + 2\,dz \\ &=^{wolfram} 1.68377. \end{aligned}$$

   Technically this gives us two roots, but we can reject the one greater than 2 since it is outside the support.

2. 2

3. 3

   The power of our test $1 - \beta$ is the probability that we reject ouir null hypothesis correctly;

$$1 - \beta = P(RH_0|H_0 \text{ is false}) = P(x_1x_2 > 0.75|\theta = 2).$$

   To find this probability we must find the joint distribution of x_1x_2. Since our variables are independent, our joint distribution is given by $f_{x_1}f_{x_2} = 4x_1x_2$, and integrating our support,

$$\int_{\frac{3}{4}}^{1} \int_{\frac{3}{4x_2}}^{1} 4x_1x_2\,dx_1\,dx_2 =^{wolfram} 0.113857$$

4. 4

5. 5

   We load our data into R:

```
> june=c(153.3, 155.9, 176.2, 189.9, 200.0, 214.9, 229.9, 231.5, 257.9, 299.9)
> december=c(151.1, 154.2, 169.9, 169.9, 185.9, 199.5, 229.9, 232.9, 279.9, 289.9)
```

   And we can use `t.test` in R since our data is normally distributed. Our null hypothesis is that the mean selling price of a condo is the same in December as in June, and our alternative is that the mean in December is less than June.

   Before using `t.test` we must check the ratio of variances, which we can do using `var.test`

```
> var.test(june,december)
```

   And the interval that this gives us does not contain 1, so we cannot conclude equal variance. So our `t.test`:

```
> t.test(december,june,alternative="less",var.equal=FALSE)
```

This gives us a p-value of 0.5841, too high to reject the null hypothesis. Therefore there is not significant evidence to suggest that the mean price of a condo decreased from December to June.

6. 6

7. 7

   Let $\hat{p}_1$ be the proportion of defects while the water was contaminated. We know this to be 16 out of 414 births. Let $\hat{p}_2$ be the defects after the water was contaminated. We know this to be 2 out of 228 births. Then our hypotheses are $H_0 : p_1 - p_2 \leq 0$, and $H_a : p_1 - p_2 > 0$. Our test statistic:

$$Z_{calc} = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim Normal(0,1)$$

   Where $\hat{p}$ is our pooled sample proportion, when simplified with our above values gives us a test statistic of 2.194423. To find our p-value, the probability of observing a sample more extreme than this, is given by:

$$P(Z > \text{Test Stat}) = \texttt{1-pnorm(2.194423)} = 0.0141.$$

   Which is less than our tolerance for type 1 error, so we conclude that the contaminated water had a significant effect on the proportion of birth defects.

8. 8

9. 9 Begin with neyman-pearson, using both sets of RVs,

$$\begin{aligned}\frac{L(H_0)}{L(H_a)} &= \frac{\prod_{i=1}^{n} 2^{y_i} e^{-2}/y_i! \prod_{j=1}^{m} 2^{x_i} e^{-2}/x_i!}{\prod_{i=1}^{n} 1/2^{y_i} e^{-\frac{1}{2}}/y_i! \prod_{j=1}^{m} 3^{x_i} e^{-3}/x_i!} \\ &= \frac{\prod_{i=1}^{n} 2^{y_i} e^{-2} \prod_{j=1}^{m} 2^{x_i} e^{-2}}{\prod_{i=1}^{n} 1/2^{y_i} e^{-\frac{1}{2}} \prod_{j=1}^{m} 3^{x_i} e^{-3}}\end{aligned}$$

$$2^{\sum y_i} \left(\frac{1}{2}\right)^{-\sum y_i} \left(\frac{2}{3}\right)^{\sum x_i} e^{-2} e^{-2} e^{\frac{1}{2}} e^{3} < k$$

$$\left(\frac{2}{3}\right)^{m\bar{x}} 4^{n\bar{y}} < k'.$$

   Taking the ln of both sides brings our variables out of the exponent but complicates the expression so I left this the way it was.

   Our test statistic is

$$\left(\frac{2}{3}\right)^{m\bar{x}} 4^{n\bar{y}}.$$

10. 10

11. 11

We use Neyman-Pearson; beginning with the ratio of our likelihood functions:

$$\begin{aligned}
\frac{L(\sigma^2=2)}{L(\sigma^2=3)} &= \frac{\prod_{i=1}^{8}\frac{1}{\sqrt{4\pi}}\exp\left[-\left(\frac{1}{4}(y_i-\mu)^2\right)\right]}{\prod_{i=1}^{8}\frac{1}{\sqrt{6\pi}}\exp\left[-\left(\frac{1}{6}(y_i-\mu)^2\right)\right]}\\
&= \frac{\frac{1}{(4\pi)^4}\prod_{i=1}^{8}\exp\left[-\left(\frac{1}{4}(y_i-\mu)^2\right)\right]}{\frac{1}{(6\pi)^4}\prod_{i=1}^{8}\exp\left[-\left(\frac{1}{6}(y_i-\mu)^2\right)\right]}\\
&= \left(\frac{6\pi}{4\pi}\right)^4\frac{\exp\left[-\frac{1}{4}\sum_{i=1}^{8}(y_i-\mu)^2\right]}{\exp\left[-\frac{1}{6}\sum_{i=1}^{8}(y_i-\mu)^2\right]}
\end{aligned}$$

$$\begin{aligned}
\left(\frac{6\pi}{4\pi}\right)^4\frac{\exp\left[-\frac{1}{4}\sum_{i=1}^{8}(y_i-\mu)^2\right]}{\exp\left[-\frac{1}{6}\sum_{i=1}^{8}(y_i-\mu)^2\right]} &< k\\
\frac{\exp\left[-\frac{1}{4}(8-1)S^2\right]}{\exp\left[-\frac{1}{6}(8-1)S^2\right]} &< k'\\
\exp\left[\frac{7}{6}S^2-\frac{7}{4}S^2\right] &< k'\\
\frac{7}{6}S^2-\frac{7}{4}S^2 &< k''\\
-\frac{14}{24}S^2 &< k''\\
S^2 &> k'''.
\end{aligned}$$

So we have our decision rule, to reject H_0 if S^2 is greater than our critical value. Now to find our critical value, setting the probability of Type I error to 0.05:

$$\begin{aligned}
0.05 &= P(RH_0|H_0 \text{ is true})\\
&= P(S^2 > CV|\sigma^2=2)\\
&= P\left(\frac{7S^2}{2} > \frac{7}{2}CV\right)\\
&= P\left(\chi^2_7 > \frac{7}{2}CV\right)
\end{aligned}$$

.

And so in R our CV is given by `qchisq(.05,7)*(2/7)`=0.6192428. Our uniformly most powerful test is to reject the null hypothesis when our sample variance is greater than this value.