Problem Sheet 7

Statistics 2

Dom Hutchinson

Question 1

a)

By assuming that the colour each bulb comes up is independent we can define $X \sim \text{Binomial}(12, p)$ to represent the number of blubs which come up red.
Define simple hypotheses $H_0 : p = 0.25$ (*i.e.* *ii*) is true) and $H_1 : p = 0.6$ (*i.e.* *i*) is true)
Define test statisic $T(X) = X$.
Since the farmer decides to accept *i*) if 8 or more bulbs come up red, we have critical value $c = 8$ and critical region $R = [8, 12]$.

b)

Power Function

$$\pi(\theta; T, c) = \pi(p; X, 8) = \mathbb{P}(X \geq 8; p) = \sum_{i=8}^{12} \binom{12}{i} p^i (1-p)^{12-i}$$

Significance Level

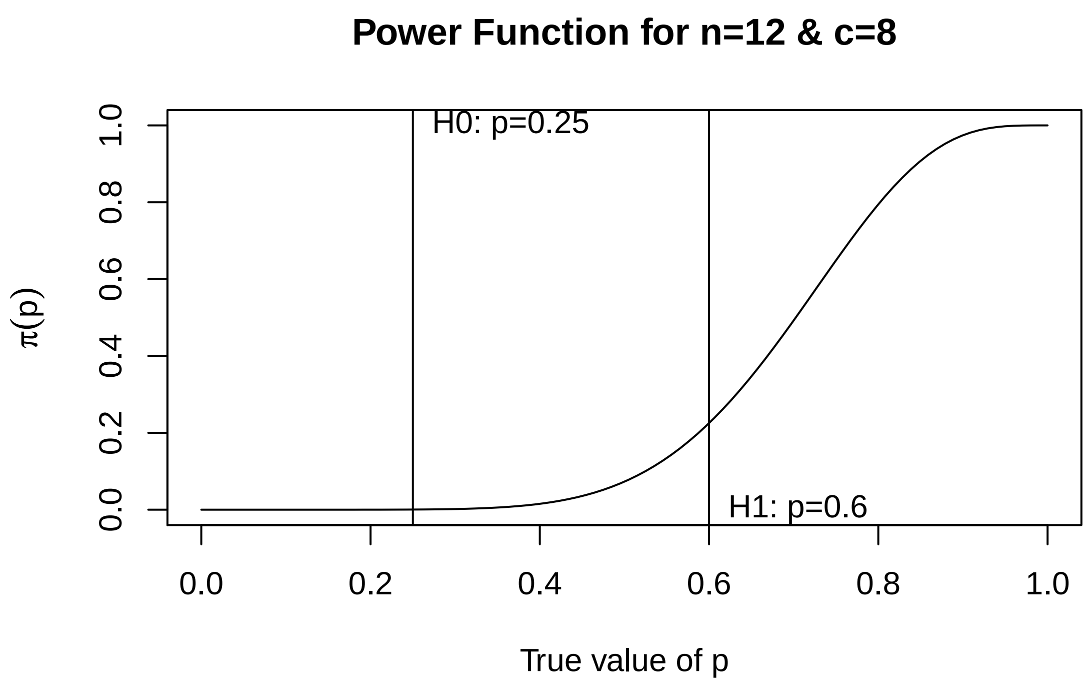
$$\alpha = \mathbb{P}(\text{Type I Error}) = \pi(0.25) = 0.0027815$$

Type II Error Level

$$\beta = \mathbb{P}(\text{Type II Error}) = 1 - \pi(0.6) = 0.7746627$$

c)

```
x<-seq(0,1,length.out=100)
y<-sapply(x,function(p) 1-pbinom(8,12,p))
plot(x,y,type="l",xlab="True value of p",ylab=expression(pi(p)),main="Power Function for n=12 & c=8")
abline(v=.25)
text(0.25,1,labels="H0: p=0.25",pos=4)
abline(v=.6)
text(0.6,0,labels="H1: p=0.6",pos=4)
```



d)

Define $T_{NP}(x)$ to be the Neyman-Pearson test. Then

$$\begin{aligned}
T_{NP}(x) &= \frac{L(p=0.6;x)}{L(p=0.25;x)} \\
&= \frac{f_X(x;p=0.6)}{f_X(x;p=0.25)} \\
&= \frac{\binom{12}{x}(0.6)^x(0.4)^{12-x}}{\binom{12}{x}(0.25)^x(0.75)^{12-x}} \\
&= \left(\frac{12}{5}\right)^x\left(\frac{8}{15}\right)^{12-x} \\
&= \frac{12^x 8^{12-x}}{5^x 15^{12-x}} \\
&= \frac{2^{2x}3^x 2^{3(12-x)}}{5^x 3^{12-x} 5^{12-x}} \\
&= \frac{2^{36-x}3^{2x-12}}{5^{12}}
\end{aligned}$$

Thus $T_{NP}(x)$ is an increasing function with x.
In **a)** we defined $T(X) = X$ meaning $T(X)$ is an equivalent test statistic to $T_{NP}(X)$. Meaning the farmer's test statistic is optimal in the Neyman-Pearson sense.

e)

Consider the power function of (T, c) with n & c not fixed. Then

$$\pi_n(p; X, c) := \sum_{i=c}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

We want a test (T, c) with signifiance level $\alpha = 0.05$ and rate of type II error $\beta = 0.1$.
Then we want to find n & c such that both the following equalities are satisfied

$$\pi_n(0.25; X, c) = 0.05 \quad \text{and} \quad 1 - \pi_n(0.6; X, c) = 0.1$$

Noting that

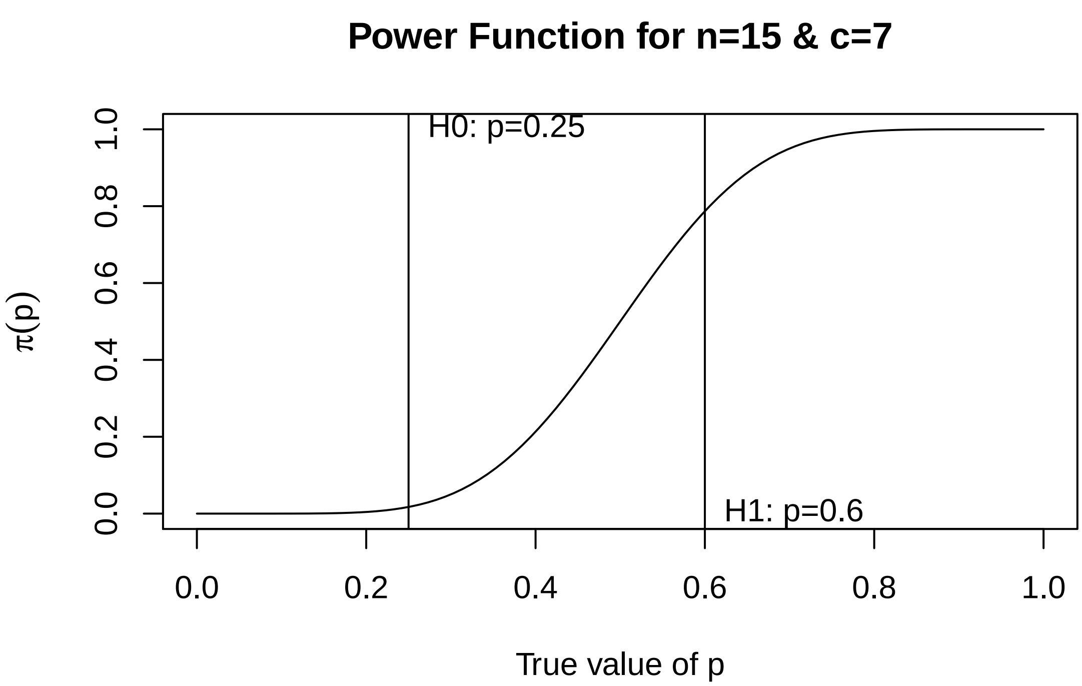
$$\pi_n(0.25; X, c) = \sum_{i=c}^{n} \binom{n}{i} \left(\frac{1}{4}\right)^i \left(\frac{3}{4}\right)^{n-i} \quad \text{and} \quad \pi_n(0.6; X, c) = \sum_{i=c}^{n} \binom{n}{i} \left(\frac{3}{5}\right)^i \left(\frac{2}{5}\right)^{n-i}$$

Using trial-and-error

n	c	$\pi_n(0.25)$	$1 - \pi_n(0.6)$
15	4	0.002	0.539
15	5	0.009	0.314
15	6	0.034	0.148
15	7	0.057	0.095

For $n = 15$ & $c = 7$ we have significance level $\alpha = 0.057$ & rate of type-II-error $\beta = 0.095$ which are both within 1 percentage point of out targets of 0.05 & 0.1 respectively.

```r
x<-seq(0,1,length.out=100)
y<-sapply(x,function(p) 1-pbinom(7,15,p))
plot(x,y,type="l",xlab="True value of p",ylab=expression(pi(p)),main="Power Function for n=15 & c=7")
abline(v=.25)
text(0.25,1,labels="H0: p=0.25",pos=4)
abline(v=.6)
text(0.6,0,labels="H1: p=0.6",pos=4)
```



Question 2

Consider a test between two simple hypotheses. For each of the following statistical models, derive the Neyman-Pearson optimal test statistic, and try to find the simplest equivalent representation.

a)

Let $\mathbf{X} \overset{\text{iid}}{\sim} \text{Poisson}(\lambda)$ and consider the tests $H_0 : \lambda = \lambda_0$ & $H_1 : \lambda = \lambda_1$ with $0 < \lambda_1 < \lambda_0$.Then

$$\begin{aligned}
T_{NP}(\mathbf{x}) &= \prod_{i=1}^{n} \frac{f_X(x_i;\lambda_1)}{f_X(x_i;\lambda_0)} \\
&= \prod_{i=1}^{n} \frac{\frac{e^{-\lambda}\lambda_1^{x_i}}{x_i!}}{\frac{e^{-\lambda}\lambda_2^{x_i}}{x_i!}} \\
&= \prod_{i=1}^{n} e^{\lambda_0-\lambda_1}\left(\frac{\lambda_1}{\lambda_0}\right)^{x_i} \\
&= e^{n(\lambda_0-\lambda_1)}\left(\frac{\lambda_1}{\lambda_0}\right)^{\sum x_i}
\end{aligned}$$

Since $\lambda_1 < \lambda_0$ then $\frac{\lambda_1}{\lambda_0} < 0$ then $T_{NP}(\mathbf{x})$ is increasing with $-S_n(\mathbf{x}) := -\sum x_i.\backslash$ Meaning $T(\mathbf{x}) := -S_n(\mathbf{x})$ is an equivalent statistic to T_{NP}.

b)

Let $\mathbf{X} \overset{\text{iid}}{\sim} \text{Exponential}(\lambda)$ and consider the tests $H_0 : \lambda = \lambda_0$ & $H_1 : \lambda = \lambda_1$ with $0 < \lambda_0 < \lambda_1$.Then

$$\begin{aligned}
T_{NP}(\mathbf{x}) &= \prod_{i=1}^{n} \frac{\mathbb{1}\{x \geq 0\}\lambda_1 e^{-\lambda_1 x_i}}{\mathbb{1}\{x \geq 0\}\lambda_0 e^{-\lambda_0 x_i}} \\
&= \frac{\mathbb{1}\{x \geq 0\}}{\mathbb{1}\{x \geq 0\}} \prod_{i=1}^{n} \frac{\lambda_1 e^{-\lambda_1 x_i}}{\lambda_0 e^{-\lambda_0 x_i}} \\
&= \frac{\mathbb{1}\{x \geq 0\}}{\mathbb{1}\{x \geq 0\}} \left(\frac{\lambda_1}{\lambda_0}\right) \prod_{i=1}^{n} e^{x_i(\lambda_0-\lambda_1)} \\
&= \frac{\mathbb{1}\{x \geq 0\}}{\mathbb{1}\{x \geq 0\}} \left(\frac{\lambda_1}{\lambda_0}\right) e^{(\lambda_0-\lambda_1)\sum x_i}
\end{aligned}$$

Note that T_{NP} is undefined if $\exists\, x_i \in \mathbf{x}$ st $x_i < 0$.
Otherwise, since $\lambda_0 < \lambda_1 \implies \lambda_0 - \lambda_1 < 0$ meaning T_{NP} is increasing with $-S_n(\mathbf{x}) := \sum x_i$.
Thus, $T(\mathbf{x}) := -S_n(\mathbf{x})$ is an equivalent statistic to $T_{NP}(\mathbf{x})$.

c)

Let $\mathbf{X} \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ with σ^2 known and consider the tests $H_0 : \mu = \mu_0$ & $H_1 : \mu = \mu_1$ with $\mu_0 < \mu_1$. Then

$$\begin{aligned}
T_{NP}(\mathbf{x}) &= \prod_{i=1}^{n} \frac{e^{-\frac{(x_i-\mu_1)^2}{2\sigma^2}}}{e^{-\frac{(x_i-\mu_0)^2}{2\sigma^2}}} \\
&= \prod_{i=1}^{n} e^{-\frac{1}{2\sigma^2}\left[(x_i-\mu_0)^2-(x_i-\mu_1)^2\right]} \\
&= e^{-\frac{1}{2\sigma^2}\sum[(x_i-\mu_0)^2-(x_i-\mu_1)^2]} \\
&= e^{-\frac{1}{2\sigma^2}\sum[2x_i(\mu_1-\mu_0)+\mu_0^2-\mu_1^2]} \\
&= e^{-\frac{1}{2\sigma^2}\left[n\mu_0^2-n\mu_1^2+2(\mu_1-\mu_0)\sum x_i\right]}
\end{aligned}$$

By the constraints we know that $\mu_1 - \mu_0 > 0$ meaning T_{NP} is increasing with $-S_n(\mathbf{x}) := \sum x_i$. Thus, $T(\mathbf{x}) := -S_n(\mathbf{x})$ is an equivalent statistic to $T_{NP}(\mathbf{x})$.

d)

Let $\mathbf{X} \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ with μ known and consider the tests $H_0 : \sigma^2 = \sigma_0^2$ & $H_1 : \sigma^2 = \sigma_1^2$ with $0 < \sigma_0^2 < \sigma_1^2$. Then

$$\begin{aligned} T_{NP}(\mathbf{x}) &= \prod_{i=1}^{n} \frac{e^{-\frac{(x_i-\mu)^2}{2\sigma_1^2}}}{e^{-\frac{(x_i-\mu)^2}{2\sigma_0^2}}} \\ &= \prod_{i=1}^{n} e^{-(x_i-\mu)^2\left[\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right]} \\ &= e^{-\left[\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}\right]\sum (x_i-\mu)^2} \end{aligned}$$

Since $\sigma_0^2 < \sigma_1^2 \implies 0 < \frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2}$ Thus T_{NP} is increasing with $-\sum(x_i - \mu)^2$ meaning $T(\mathbf{x}) = -\sum(x_i - \mu)^2$ is an equivalent statistic to T_{NP}.