STAT 410 Fall 2020

Homework #11

A. Stepanov

(due Friday, December 4, by 5:00 p.m. CST)

No credit will be given without supporting work.

11. Suppose that students arrive to a certain on-campus COVID-19 testing location according to a Poisson process with average rate λ per minute. We plan to observe this location for the next 4 minutes in order to test $H_0: \lambda = 3$ vs. $H_1: \lambda < 3$.

That is, let X_1, X_2, X_3, X_4 be a random sample of size $n = 4$ from a Poisson distribution with mean λ. Consider the test $H_0: \lambda = 3$ vs. $H_1: \lambda < 3$.

a) Find the best rejection region with the significance level closest to 0.05.

Reject H_0 if $X_1 + X_2 + X_3 + X_4 \le c$.

$X_1 + X_2 + X_3 + X_4$ has Poisson distribution with mean 4λ.

$$\alpha = P(\text{Reject } H_0 \mid H_0 \text{ is true}) = P(X_1 + X_2 + X_3 + X_4 \le c \mid \lambda = 3)$$

$$= P(\text{Poisson}(4 \cdot 3) \le c) = P(\text{Poisson}(12) \le c).$$

$P(\text{Poisson}(12) \le 6) = 0.046 \approx 0.05.$

Reject H_0 if $\mathbf{X_1 + X_2 + X_3 + X_4 \le 6}$.

b) What is the power of the Rejection Region obtained in part (a) if $\lambda = 2$? if $\lambda = 1.2$?

$$\text{Power}(\lambda = 2) = P(X_1 + X_2 + X_3 + X_4 \le 6 \mid \lambda = 2) = P(\text{Poisson}(4\lambda) \le 6 \mid \lambda = 2)$$

$$= P(\text{Poisson}(8) \le 6) = \mathbf{0.313}.$$

$$\text{Power}(\lambda = 1.2) = P(X_1 + X_2 + X_3 + X_4 \le 6 \mid \lambda = 1.2)$$

$$= P(\text{Poisson}(4\lambda) \le 6 \mid \lambda = 1.2)$$

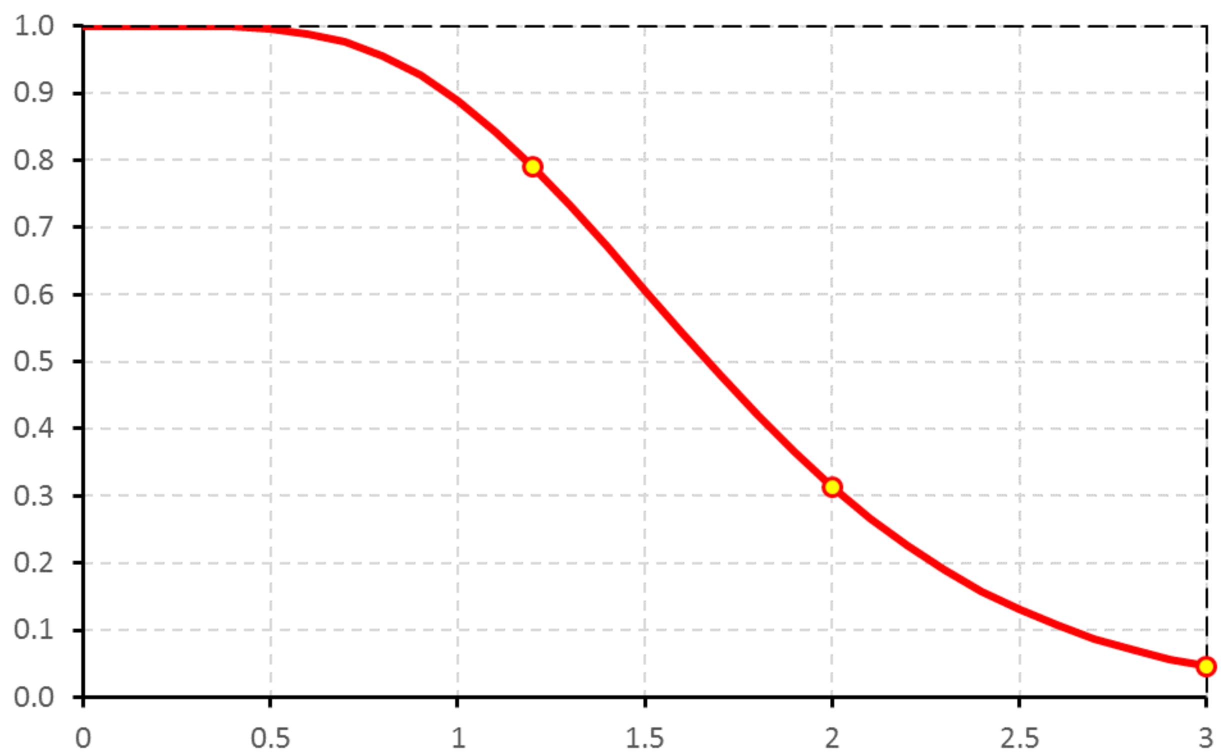
$$= P(\text{Poisson}(4.8) \le 6) = \mathbf{0.791}.$$

c) Suppose $x_1 = 1,\ x_2 = 3,\ x_3 = 1,\ x_4 = 2$. Find the p-value of the test.

$$x_1 + x_2 + x_3 + x_4 = 7.$$

$$\text{P-value} = P(\text{value of } \textstyle\sum_{i=1}^{n=4} X_i \text{ as extreme or more extreme than } 7 \mid H_0 \text{ true})$$

$$= P(X_1 + X_2 + X_3 + X_4 \le 7 \mid \lambda = 3) = P(\text{Poisson}(12) \le 7) = \mathbf{0.090}.$$



7. Let $\psi > 0$ be a population parameter, and let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with probability density function

$$f(x;\psi) = \frac{2}{\sqrt{\pi\psi}}\, e^{-x^2/\psi}, \qquad x > 0, \qquad \text{zero otherwise.}$$

Recall: $W = X^2$ has Gamma$(\alpha = \frac{1}{2}, \theta = \psi)$ distribution.

$\sum_{i=1}^{n} X_i^2$ is a sufficient statistic for ψ.

Suppose $n = 4$. We wish to test $H_0: \psi = 1.5$ vs. $H_1: \psi > 1.5$.

p) Consider rejection region Reject H_0 if $\sum_{i=1}^{n=4} x_i^2 \ge 7.2$. Find …

i) … the significance level α;

ii) … the power if $\psi = 2$ and if $\psi = 4$.

$\sum_{i=1}^{n=4} X_i^2$ has Gamma$(\alpha = \frac{n}{2} = 2, \theta = \psi)$ distribution.

(i) α = significance level = $P(\text{Reject } H_0 \mid H_0 \text{ is true}) = P(\sum_{i=1}^{n=4} X_i^2 \ge 7.2 \mid \psi = 1.5)$

$= P(\text{Gamma}(\alpha = 2,\ \theta = 1.5) \ge 7.2)$

```
> 1-pgamma(7.2,2,1/1.5)
[1] 0.04773253
```

$= P(\text{Poisson}(\frac{7.2}{1.5}) \le 2 - 1) = P(\text{Poisson}(4.8) \le 1) =$ **0.048**.

```
> ppois(2-1,7.2/1.5)
[1] 0.04773253
```

$$= P\left(\frac{2}{1.5} T_2 \geq \frac{2}{1.5} \cdot 7.2 \,\middle|\, \theta = 1.5 \right) = P(\chi^2(4) \geq 9.6).$$

```
> 1-pchisq(9.6,4)
[1] 0.04773253
```

$$\int_{7.2}^{\infty} \frac{1}{1.5^2} x^{2-1} e^{-\frac{x}{1.5}} dx = 0.04773\ldots$$

(ii) $$\text{Power}(\psi = 2) = P(\text{Reject } H_0 \mid \psi = 2) = P\left(\sum_{i=1}^{n=4} X_i^2 \geq 7.2 \,\middle|\, \psi = 2 \right)$$

$$= P(\text{Gamma}(\alpha = 2,\ \theta = 2) \geq 7.2)$$

```
> 1-pgamma(7.2,2,1/2)
[1] 0.1256891
```

$$= P\left(\text{Poisson}\left(\frac{7.2}{2} \right) \leq 2 - 1 \right) = P(\text{Poisson}(3.6) \leq 1) = \mathbf{0.126}.$$

```
> ppois(2-1,7.2/2)
[1] 0.1256891
```

$$= P\left(\frac{2}{2} T_2 \geq \frac{2}{2} \cdot 7.2 \,\middle|\, \theta = 2 \right) = P(\chi^2(4) \geq 7.2).$$

```
> 1-pchisq(7.2,4)
[1] 0.1256891
```

$$\int_{7.2}^{\infty} \frac{1}{2^2} x^{2-1} e^{-\frac{x}{2}} dx = 0.12568\ldots$$

$$\text{Power}(\psi = 4) = P(\text{Reject } H_0 \mid \psi = 4) = P\left(\sum_{i=1}^{n=4} X_i^2 \geq 7.2 \,\middle|\, \psi = 4 \right)$$

$$= P(\text{Gamma}(\alpha = 2,\ \theta = 4) \geq 7.2)$$

```
> 1-pgamma(7.2,2,1/4)
[1] 0.4628369
```

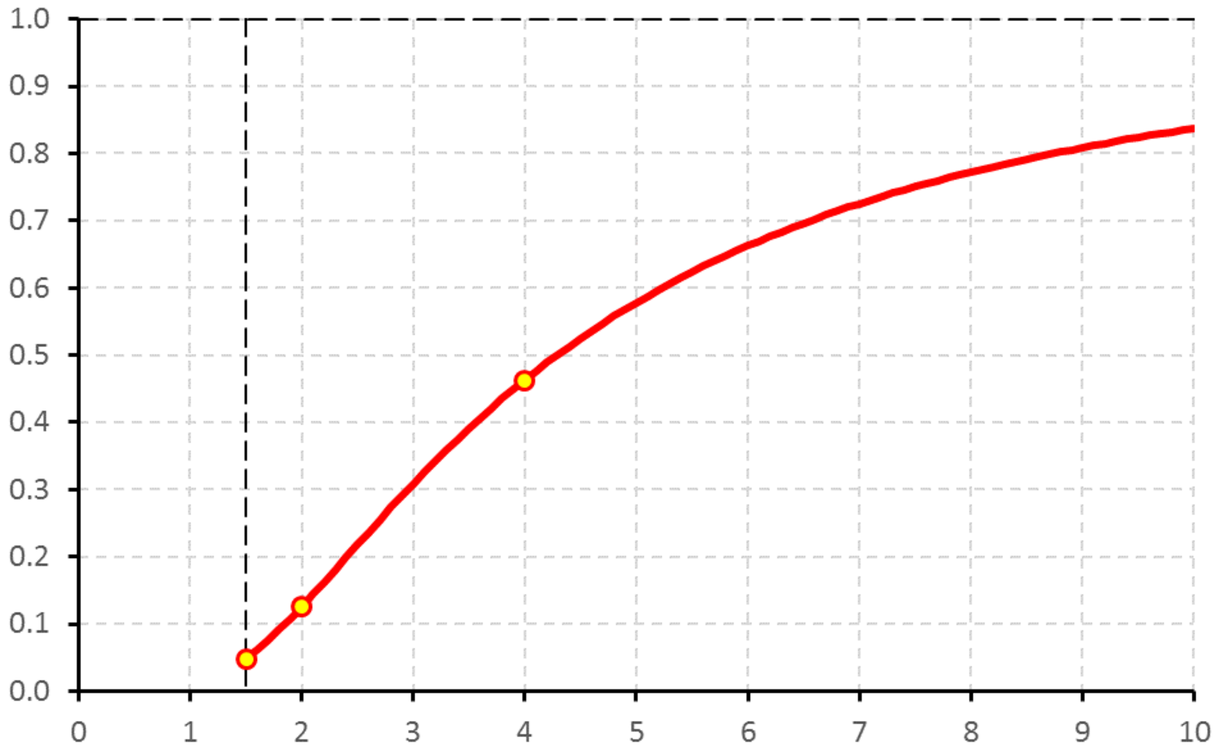
$$= \mathrm{P}\left(\mathrm{Poisson}\left(\frac{7.2}{4}\right) \le 2-1\right) = \mathrm{P}\left(\mathrm{Poisson}(1.8) \le 1\right) = \mathbf{0.463}.$$

```
> ppois(2-1,7.2/4)
[1] 0.4628369
```

$$= \mathrm{P}\left(\frac{2}{4}\,\mathrm{T}_2 \ge \frac{2}{4}\cdot 7.2 \,\middle|\, \theta = 4\right) = \mathrm{P}\left(\chi^2(4) \ge 3.6\right).$$

```
> 1-pchisq(3.6,4)
[1] 0.4628369
```

$$\int_{7.2}^{\infty} \frac{1}{4^2} x^{2-1} e^{-\frac{x}{4}} dx = 0.46283\ldots$$



q) Suppose $x_1 = 0.2, \quad x_2 = 0.6, \quad x_3 = 1.1, \quad x_4 = 1.7.$

Find the p-value of the test.

$$\sum_{i=1}^{n=4} x_i^2 = 0.2^2 + 0.6^2 + 1.1^2 + 1.7^2 = 4.5.$$

$$\text{p-value} = P\Big(\sum_{i=1}^{n=4} X_i^2 \text{ as extreme or more extreme than } \big(\sum_{i=1}^{n=4} x_i^2 \big)_{\text{observed}} \,\big|\, H_0 \text{ true} \Big)$$

$$= P\Big(\sum_{i=1}^{n=4} X_i^2 \ge 4.5 \,\big|\, \psi = 1.5 \Big)$$

$$= P\big(\text{Gamma}(\alpha = 2,\ \theta = 1.5) \ge 4.5 \big)$$

```
> 1-pgamma(4.5,2,1/1.5)
[1] 0.1991483
```

$$= P\Big(\text{Poisson}\Big(\frac{4.5}{1.5} \Big) \le 2 - 1 \Big) = P\big(\text{Poisson}(3.0) \le 1 \big) = \mathbf{0.199}.$$

```
> ppois(2-1,4.5/1.5)
[1] 0.1991483
```

$$= P\Big(\frac{2}{1.5} T_2 \ge \frac{2}{1.5} \cdot 4.5 \,\big|\, \theta = 1.5 \Big) = P\big(\chi^2(4) \ge 6.0 \big).$$

```
> 1-pchisq(6.0,4)
[1] 0.1991483
```

$$\int_{4.5}^{\infty} \frac{1}{1.5^2} x^{2-1} e^{-\frac{x}{1.5}} dx = 0.19914\ldots$$

For fun:

r) Find a uniformly most powerful rejection region with the significance level $\alpha = 0.05$.

$H_0: \psi = 1.5$ vs. $H_1: \psi > 1.5$. $n = 4$.

Let $\psi > 1.5$.

$$\frac{L(1.5)}{L(\psi)} = \frac{L(1.5;\, x_1, x_2, \ldots, x_n)}{L(\psi;\, x_1, x_2, \ldots, x_n)} = \frac{\prod_{i=1}^{n} \frac{2}{\sqrt{\pi 1.5}}\, e^{-x_i^2/1.5}}{\prod_{i=1}^{n} \frac{2}{\sqrt{\pi\psi}}\, e^{-x_i^2/\psi}}$$

$$= \left(\frac{\psi}{1.5}\right)^{n/2} \exp\left\{\left(\frac{1}{\psi} - \frac{1}{1.5}\right)\sum_{i=1}^{n} x_i^2\right\} \le k.$$

$$\Leftrightarrow \qquad \exp\left\{\left(\frac{1}{\psi} - \frac{1}{1.5}\right)\sum_{i=1}^{n} x_i^2\right\} \le k_1.$$

$$\Leftrightarrow \qquad \left(\frac{1}{\psi} - \frac{1}{1.5}\right)\sum_{i=1}^{n} x_i^2 \le k_2.$$

$$\psi > 1.5 \qquad \Rightarrow \qquad \frac{1}{\psi} - \frac{1}{1.5} < 0$$

$$\Leftrightarrow \qquad \sum_{i=1}^{n} x_i^2 \ge c.$$

Intuition: ψ is "θ". $E(W) = \alpha\,\theta = \frac{1}{2}\psi.$

Large ψ $\Rightarrow$ large $w = x^2$.

The sign is the same as the sign in H_1.

$\sum_{i=1}^{n=4} X_i^2$ has Gamma$(\alpha = \frac{n}{2} = 2, \theta = \psi)$ distribution. $\qquad \sum_{i=1}^{n=4} X_i^2 = T_2$.

$$0.05 = \alpha = P(\text{Reject } H_0 \mid H_0 \text{ is true}) = P(\sum_{i=1}^{n=4} X_i^2 \ge c \mid \psi = 1.5)$$

$$= P(T_2 \ge c \mid \theta = 1.5)$$

If T_α has a Gamma(α, θ) distribution, then $\frac{2}{\theta} T_\alpha$ has a $\chi^2(2\alpha)$ distribution.

$$= P(\frac{2}{\theta} T_2 \ge \frac{2}{\theta} c \mid \theta = 1.5) = P(\frac{2}{1.5} T_2 \ge \frac{2}{1.5} c \mid \theta = 1.5)$$

$$= P(\chi^2(4) \ge \frac{2}{1.5} c).$$

$\Rightarrow \quad \frac{2}{1.5} c = \chi^2_{0.05}(4) = 9.488. \qquad \Rightarrow \quad c = \mathbf{7.116}.$

Reject H_0 if $\sum_{i=1}^{n=4} x_i^2 \ge \mathbf{7.116}$.

```
> qgamma(0.95,2,1/1.5)
[1] 7.115797
>
> qchisq(0.95,2*2)
[1] 9.487729
> qchisq(0.95,2*2)*(1.5/2)
[1] 7.115797
```

8. Let $\beta > 0$ be a population parameter, and let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with probability density function

$$f(x;\beta) = \beta\,(1-x)^{\beta-1}, \qquad 0 < x < 1, \qquad \text{zero otherwise.}$$

Recall: $W = -\ln(1-X)$ has an $\text{Exponential}\left(\theta = \frac{1}{\beta}\right)$

$= \text{Gamma}\left(\alpha = 1, \theta = \frac{1}{\beta}\right)$ distribution.

$\sum_{i=1}^{n} \left(-\ln(1-X_i)\right)$ is a sufficient statistic for β.

Suppose $n = 3$. We wish to test $H_0: \beta = 0.2$ vs. $H_1: \beta > 0.2$.

p) Find the uniformly most powerful rejection region with significance level $\alpha = 0.10$.

Hint 1: Let $\beta > 0.2$. Start with

$$\frac{L(H_0;\, x_1, x_2, \ldots, x_n)}{L(H_1;\, x_1, x_2, \ldots, x_n)} = \frac{L(0.2;\, x_1, x_2, \ldots, x_n)}{L(\beta;\, x_1, x_2, \ldots, x_n)} = \frac{\prod_{i=1}^{n} f(x_i;\, 0.2)}{\prod_{i=1}^{n} f(x_i;\, \beta)} \le k.$$

Simplify this. Since $Y = \sum_{i=1}^{n} \left(-\ln(1-X_i)\right)$ is a sufficient statistic for β,

and the final form of the "best" rejection region should look like this:

"Reject H_0 if $\sum_{i=1}^{n} \left(-\ln(1-x_i)\right) [\le \text{or} \ge]\; c$".

The direction of the inequality sign is what you are trying to determine.

Hint 2: $Y = \sum_{i=1}^{n} \left(-\ln(1-X_i)\right) = \sum_{i=1}^{n} W_i$ has a $\text{Gamma}\left(\alpha = n,\ \theta = \frac{1}{\beta}\right)$ distribution.

Hint 3: Want c such that

$$0.10 = \alpha = P(\text{Reject } H_0 \mid H_0 \text{ is true}) = P\left(\sum_{i=1}^{n=3} \left(-\ln(1-X_i)\right) \;?\; c \mid \beta = 0.2\right).$$

Hint 4: If T_α has a $\text{Gamma}(\alpha, \theta = 1/\lambda)$ distribution,

then $\frac{2}{\theta}\mathrm{T}_\alpha = 2\lambda\mathrm{T}_\alpha$ has a $\chi^2(2\alpha)$ distribution.

Let $\beta > 0.2$.

$$\frac{\mathrm{L}(0.2)}{\mathrm{L}(\beta)} = \frac{\mathrm{L}(0.2;\, x_1, x_2, \ldots, x_n)}{\mathrm{L}(\beta;\, x_1, x_2, \ldots, x_n)} = \frac{\prod_{i=1}^{n} 0.2\,(1-x_i)^{-0.8}}{\prod_{i=1}^{n} \beta\,(1-x_i)^{\beta-1}}$$

$$= \left(\frac{0.2}{\beta}\right)^n \left(\prod_{i=1}^{n}(1-x_i)\right)^{0.2-\beta} \le k.$$

$$\Leftrightarrow \qquad \left(\prod_{i=1}^{n}(1-x_i)\right)^{0.2-\beta} \le k_1.$$

$$\Leftrightarrow \qquad (\beta-0.2)\sum_{i=1}^{n}\left(-\ln(1-x_i)\right) \le k_2.$$

$\beta > 0.2 \quad \Rightarrow \quad \beta - 0.2 > 0$

$$\Leftrightarrow \qquad \sum_{i=1}^{n}\left(-\ln(1-x_i)\right) \le c.$$

Intuition: β is "λ". $\mathrm{E}(\mathrm{W}) = \alpha\,\theta = \frac{1}{\beta}$.

Large $\beta \quad \Rightarrow \quad$ small $w = -\ln(1-x)$.

The sign is opposite from the sign in H_1.

$\sum_{i=1}^{n=3}\left(-\ln(1-\mathrm{X}_i)\right)$ has $\text{Gamma}(\alpha = n = 3, \theta = \frac{1}{\beta})$ distribution.

$$\sum_{i=1}^{n=3} \left(-\ln(1-X_i) \right) = T_3.$$

$$0.10 = \alpha = P(\text{Reject } H_0 \mid H_0 \text{ is true}) = P\left(\sum_{i=1}^{n=3} \left(-\ln(1-X_i) \right) \le c \mid \beta = 0.2 \right)$$

$$= P(T_3 \le c \mid \lambda = 0.2) = P(T_3 \le c \mid \theta = 5)$$

If T_α has a $\text{Gamma}(\alpha, \theta)$ distribution,

then $\frac{2}{\theta} T_\alpha = 2\lambda T_\alpha$ has a $\chi^2(2\alpha)$ distribution.

$$= P(2\lambda T_3 \le 2\lambda c \mid \lambda = 0.2) = P\left(\frac{2}{5} T_3 \le \frac{2}{5} c \mid \theta = 5\right)$$

$$= P(2 \cdot 0.2 \cdot T_3 \le 2 \cdot 0.2 \cdot c \mid \lambda = 0.2) = P\left(\frac{2}{5} T_3 \le \frac{2}{5} c \mid \theta = 5\right)$$

$$= P(\chi^2(6) \le 0.4\,c).$$

$\Rightarrow \qquad 0.4\,c = \chi^2_{0.90}(6) = 2.204. \qquad\qquad \Rightarrow \qquad c = \mathbf{5.51}.$

Reject H_0 if $\sum_{i=1}^{n=3} \left(-\ln(1-x_i) \right) \le \mathbf{5.51}$.

```
> qgamma(0.10,3,0.2)
[1] 5.510327
>
> qchisq(0.10,2*3)
[1] 2.204131
> qchisq(0.10,2*3)/(2*0.2)
[1] 5.510327
```

q) Suppose $x_1 = 0.31$, $x_2 = 0.77$, $x_3 = 0.93$. Find the p-value of the test.

Hint 1: Probability that $\sum_{i=1}^{n=3} \left(-\ln(1 - X_i) \right)$ is as extreme or more extreme than the observed $\sum_{i=1}^{n=3} \left(-\ln(1 - x_i) \right)$ …

Hint 2: For the p-value, go in the same direction as the "best" rejection region.

Hint 3: … computed under the assumption that H_0 is true.

$$\sum_{i=1}^{n=3} \left(-\ln(1 - x_i) \right) = -\ln 0.69 - \ln 0.23 - \ln 0.07 \approx 4.5.$$

$$\text{p-value} = P\left(\sum_{i=1}^{n=3} \left(-\ln(1 - X_i) \right) \le 4.5 \mid \beta = 0.2 \right)$$

$$= P\left(\text{Gamma}(\alpha = 3,\ \lambda = 0.2,\ \theta = 5) \le 4.5 \right)$$

```
> pgamma(4.5,3,0.2)
[1] 0.06285693
```

$$= P\left(\text{Poisson}(4.5 \cdot 0.2) \ge 3 \right) = 1 - P\left(\text{Poisson}(0.9) \le 2 \right)$$

$$= 1 - 0.937 = \mathbf{0.063}.$$

```
> 1-ppois(3-1,0.2*4.5)
[1] 0.06285693
```

$$= P\left(2 \cdot 0.2 \cdot T_3 \le 2 \cdot 0.2 \cdot 4.5 \mid \lambda = 0.2 \right) = P\left(\chi^2(6) \le 1.8 \right).$$

```
> pchisq(2*0.2*4.5,2*3)
[1] 0.06285693
```

$$\int_0^{4.5} \frac{0.2^3}{2} x^{3-1} e^{-0.2x} dx = 0.06285\ldots$$

r) (i) **Using the p-value obtained in part (q)**, state your decision (Reject H_0 or Do NOT Reject H_0) at $\alpha = 0.05$.

(ii) **Using the rejection region obtained in part (p)**, state your decision (Reject H_0 or Do NOT Reject H_0) at $\alpha = 0.10$.

(i) p-value $> \alpha \quad \Rightarrow \quad$ Do NOT Reject H_0.

p-value $< \alpha \quad \Rightarrow \quad$ Reject H_0.

Since $0.063 > 0.05$, **Do NOT Reject $\mathbf{H}_0$ at $\alpha = 0.05$.**

(ii) Reject H_0 if $\sum_{i=1}^{n=3} \left(-\ln(1 - x_i) \right) \le 5.51$.

$\sum_{i=1}^{n=3} \left(-\ln(1 - x_i) \right) \approx 4.5$. $\qquad 4.5 \le 5.51$.

Reject $\mathbf{H}_0$ at $\alpha = 0.10$.