Examples for 11/18/2020 (2)

Fall 2020
A. Stepanov

Examples for 10/19/2020 (2) and 10/23/2020 (2)
and 10/30/2020 (3) and 11/04/2020 (2) (continued)

1. Let $\beta > 0$ and let $X_1, X_2, \dots, X_n$ be a random sample from the distribution with probability density function

$$f(x;\beta) = \frac{\beta^2 \ln x}{x^{\beta+1}}, \qquad x > 1, \qquad \text{zero otherwise.}$$

Recall: Examples for 11/04/2020 (2):

$$Y = \sum_{i=1}^{n} \ln X_i \quad \text{is a sufficient statistic for } \beta.$$

Examples for 10/19/2020 (2):

$$W = \ln X \text{ has a Gamma}(\alpha = 2, \theta = \frac{1}{\beta}) \text{ distribution.}$$

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

We wish to test $H_0: \beta = 1.4$ vs. $H_1: \beta > 1.4$.

a) Find the form of the uniformly most powerful rejection region.

Hint: Let $\beta > 1.4$. Start with

$$\frac{L(H_0;\, x_1, x_2, \dots, x_n)}{L(H_1;\, x_1, x_2, \dots, x_n)} = \frac{L(1.4;\, x_1, x_2, \dots, x_n)}{L(\beta;\, x_1, x_2, \dots, x_n)} = \frac{\prod_{i=1}^{n} f(x_i; 1.4)}{\prod_{i=1}^{n} f(x_i; \beta)} \le k.$$

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

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

"Reject H_0 if $\sum_{i=1}^{n} \ln x_i \, [\le \text{ or } \ge] \, c$".

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

Let $\beta > 1.4$.

$$\frac{L(1.4)}{L(\beta)} = \frac{\prod_{i=1}^{n} \frac{1.4^2 \ln x_i}{x_i^{1.4+1}}}{\prod_{i=1}^{n} \frac{\beta^2 \ln x_i}{x_i^{\beta+1}}} = \left(\frac{1.4}{\beta}\right)^{2n} \left(\prod_{i=1}^{n} x_i\right)^{\beta-1.4} \le k.$$

$$\Leftrightarrow \qquad \left(\prod_{i=1}^{n} x_i\right)^{\beta-1.4} \le k_1 = k\left(\frac{\beta}{1.4}\right)^{2n}.$$

$$\Leftrightarrow \qquad (\beta - 1.4) \sum_{i=1}^{n} \ln x_i \le k_2 = \ln k_1.$$

$$\Leftrightarrow \qquad \sum_{i=1}^{n} \ln x_i \le c = \frac{k_2}{\beta - 1.4}, \qquad \text{since } \beta > 1.4.$$

Intuition: β is "λ". $E(W) = \alpha\,\theta = \dfrac{2}{\beta}$.

Large β $\Rightarrow$ small $\ln x$.

The sign is opposite from the sign in H_1.

Reject H_0 if $\sum_{i=1}^{n} \ln x_i \le c$.

b) Suppose $n = 5$. Find the uniformly most powerful rejection region with $\alpha = 0.05$.

Hint 1: $\mathrm{Y} = \sum_{i=1}^{n} \ln \mathrm{X}_i = \sum_{i=1}^{n} \mathrm{W}_i$ has a Gamma$(\alpha = 2n,\ \theta = \frac{1}{\beta})$ distribution.

Hint 2: Want c such that

$$0.05 = \alpha = \mathrm{P}(\text{Reject } \mathrm{H}_0 \mid \mathrm{H}_0 \text{ is true}) = \mathrm{P}(\sum_{i=1}^{5} \ln \mathrm{X}_i \ ?\ c \mid \beta = 1.4).$$

Hint 3: If T has a Gamma$(\alpha, \theta = {}^{1}\!/_{\lambda})$ distribution, then

${}^{2\mathrm{T}}\!/_{\theta} = 2\lambda\mathrm{T}$ has a $\chi^2(2\alpha)$ distribution (a chi-square distribution with 2α degrees of freedom).

$\sum_{i=1}^{n} \ln \mathrm{X}_i = \sum_{i=1}^{n} \mathrm{W}_i$ has a Gamma$(\alpha = 2n = 10,\ \theta = \frac{1}{\beta})$ distribution.

Then $2\beta \sum_{i=1}^{5} \ln \mathrm{X}_i$ has a $\chi^2(2\alpha = 4n = 20$ degrees of freedom$)$ distribution.

$$0.05 = \mathrm{P}(\sum_{i=1}^{5} \ln \mathrm{X}_i \le c \mid \beta = 1.4) = \mathrm{P}(2\beta \sum_{i=1}^{5} \ln \mathrm{X}_i \le 2\beta c \mid \beta = 1.4)$$

$$= \mathrm{P}(\chi^2(20) \le 2.8\,c).$$

$\Rightarrow \quad 2.8\,c = \chi^2_{0.95}(20) = 10.85. \qquad \Rightarrow \qquad c = \mathbf{3.875}.$

The uniformly most powerful rejection region is "Reject H_0 if $\sum_{i=1}^{5} \ln x_i \le 3.875$."

So $\alpha = 0.05$… Yes.

And $\alpha = 10$… Yes. ☹

c) Find the power of the rejection region from (b) if $\beta = 4$.

Hint: $\text{Power}(\beta) = P(\text{Reject } H_0 \mid H_0 \text{ is NOT true}) = P(\sum_{i=1}^{5} \ln X_i \; ? \; c \mid \beta = 4)$.

Suggestion: If T has a $\text{Gamma}(\alpha, \theta = 1/\lambda)$ distribution, where α is an integer,

then $F_T(t) = P(T \le t) = P(X_t \ge \alpha)$ and $P(T > t) = P(X_t \le \alpha - 1)$,

where X_t has a $\text{Poisson}(\lambda t = t/\theta)$ distribution.

$$\text{Power}(\beta = 4) = P(\sum_{i=1}^{5} \ln X_i \le 3.875 \mid \beta = 4)$$

$$= P(\text{Gamma}(\alpha = 10,\ \theta = \frac{1}{4}) \le 3.875)$$

$$= P(\text{Poisson}(3.875 \cdot 4) \ge 10) = 1 - P(\text{Poisson}(15.5) \le 9)$$

$$= 1 - 0.055 = \mathbf{0.945}.$$

```
> 1-ppois(9,15.5)
[1] 0.9448095
```

$$\text{Power}(\beta = 4) = P(\sum_{i=1}^{5} \ln X_i \le 3.875 \mid \beta = 4)$$

$$= P(\text{Gamma}(\alpha = 10,\ \theta = \frac{1}{4}) \le 3.875)$$

$$= P(\chi^2(20) \le 8 \cdot 3.875) = P(\chi^2(20) \le 31).$$

```
> pchisq(31,20)
[1] 0.9448095
```

```
> pgamma(3.875,10,4)
[1] 0.9448095
```

d) Find the power of the rejection region from (b) if $\beta = 2$.

Hint 1: $\text{Power}(\beta) = P(\text{Reject } H_0 \mid H_0 \text{ is NOT true}) = P(\sum_{i=1}^{5} \ln X_i \ ?\ c \mid \beta = 2)$.

Hint 2: Excel: =GAMMA.DIST($x, \alpha, \theta, 1$) $P(\text{Gamma}(\alpha, \theta) \le x)$,

R: `pgamma(` x `,` α `,` λ `)` $P(\text{Gamma}(\alpha, 1/\lambda) \le x)$.

OR

If T has a $\text{Gamma}(\alpha, \theta = 1/\lambda)$ distribution, where α is an integer,

then $F_T(t) = P(T \le t) = P(X_t \ge \alpha)$ and $P(T > t) = P(X_t \le \alpha - 1)$,

where X_t has a $\text{Poisson}(\lambda t = t/\theta)$ distribution.

Excel: =POISSON.DIST($x, \lambda, 1$) $P(\text{Poisson}(\lambda) \le x)$,

=POISSON.DIST($x, \lambda, 0$) $P(\text{Poisson}(\lambda) = x)$.

R: `ppois(` x `,` λ `)` $P(\text{Poisson}(\lambda) \le x)$,

`dpois(` x `,` λ `)` $P(\text{Poisson}(\lambda) = x)$.

OR

If T has a $\text{Gamma}(\alpha, \theta = 1/\lambda)$ distribution, then

$2T/\theta = 2\lambda T$ has a $\chi^2(2\alpha)$ distribution (a chi-square distribution with 2α degrees of freedom).

Excel: =CHISQ.DIST(x, degrees of freedom, 1) $P(\chi^2(\text{d.f.}) \le x)$,

=CHISQ.DIST.RT(x, degrees of freedom) $P(\chi^2(\text{d.f.}) > x)$.

R: `pchisq(` x `,` degrees of freedom `)` $P(\chi^2(\text{d.f.}) \le x)$,

`pchisq(` x `,` degrees of freedom `, lower.tail=FALSE)`.

$$\text{Power}(\beta = 2) = P\left(\sum_{i=1}^{5} \ln X_i \le 3.875 \,\middle|\, \beta = 2\right)$$

$$= P\left(\text{Gamma}\left(\alpha = 10,\ \theta = \frac{1}{2}\right) \le 3.875\right)$$

$$= P\left(\text{Poisson}(3.875 \cdot 2) \ge 10\right) = 1 - P\left(\text{Poisson}(7.75) \le 9\right)$$

```
> 1-ppois(9,7.75)
[1] 0.2528812
```

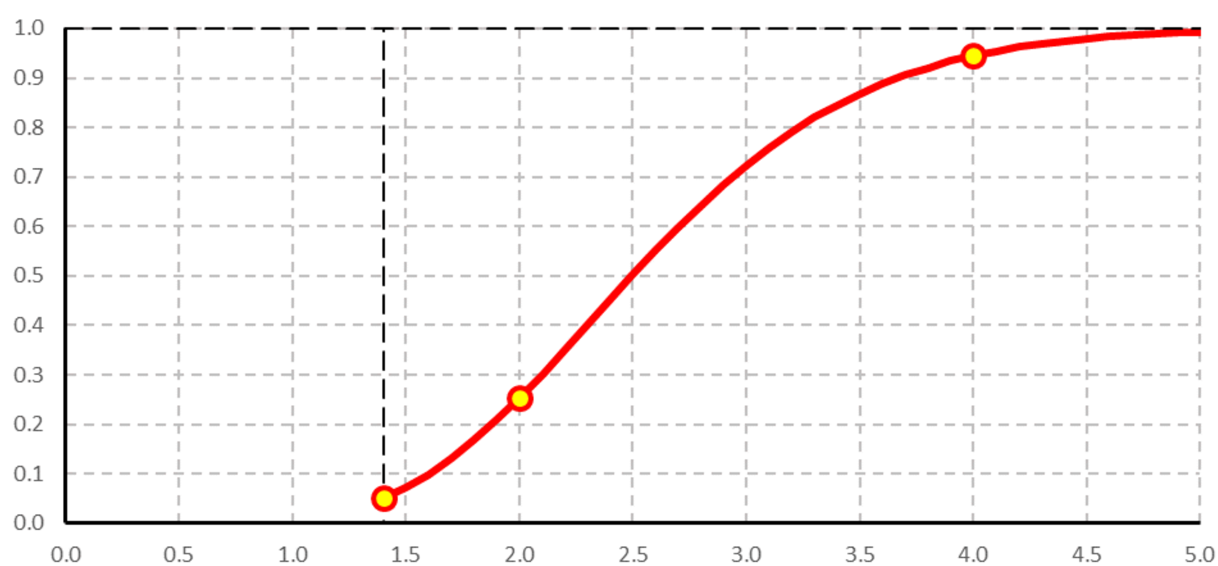
$$\text{Power}(\beta = 4) = P\left(\sum_{i=1}^{5} \ln X_i \le 3.875 \,\middle|\, \beta = 2\right)$$

$$= P\left(\text{Gamma}\left(\alpha = 10,\ \theta = \frac{1}{2}\right) \le 3.875\right)$$

$$= P\left(\chi^2(20) \le 4 \cdot 3.875\right) = P\left(\chi^2(20) \le 15.5\right).$$

```
> pchisq(15.5,20)
[1] 0.2528812
```

```
> pgamma(3.875,10,2)
[1] 0.2528812
```



1. (continued)

Suppose $n = 5$, and $x_1 = 1.3$, $x_2 = 1.4$, $x_3 = 2.0$, $x_4 = 3.0$, $x_5 = 5.0$.

e) Find the p-value for the test.

Hint 1: $\ldots \sum_{i=1}^{5} \ln X_i$ as extreme or more extreme than the observed $\sum_{i=1}^{n} \ln x_i \ldots$

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.

Suggestion: If T has a $\text{Gamma}(\alpha, \theta = 1/\lambda)$ distribution, where α is an integer,

then $F_T(t) = P(T \le t) = P(X_t \ge \alpha)$ and $P(T > t) = P(X_t \le \alpha - 1)$,

where X_t has a $\text{Poisson}(\lambda t = t/\theta)$ distribution.

$$\sum_{i=1}^{5} \ln x_i = \ln 1.3 + \ln 1.4 + \ln 2.0 + \ln 3.0 + \ln 5.0 \approx 4.$$

$$\text{P-value} = P\left(\sum_{i=1}^{5} \ln X_i \le 4 \,\middle|\, \beta = 1.4\right) = P\left(\text{Gamma}\left(\alpha = 10,\ \theta = \frac{1}{1.4}\right) \le 4\right)$$

$$= P(\text{Poisson}(4 \cdot 1.4) \ge 10) = 1 - P(\text{Poisson}(5.6) \le 9)$$

$$= 1 - 0.941 = \mathbf{0.059}.$$

```
> 1-ppois(9,5.6)
[1] 0.05912997
```

```
> pgamma(4,10,1.4)
[1] 0.05912997
```

$$\text{P-value} = P\left(\sum_{i=1}^{5} \ln X_i \le 4 \;\middle|\; \beta = 1.4 \right) = P\left(\text{Gamma}\left(\alpha = 10,\ \theta = \frac{1}{1.4} \right) \le 4 \right)$$

$$= P\left(\chi^2(20) \le 2.8 \cdot 4 \right) = P\left(\chi^2(20) \le 11.2 \right).$$

```
> pchisq(11.2,20)
[1] 0.05912997
```

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

(ii) Does your decision agree with part (b)? That is, does the observed value of $\sum_{i=1}^{n} \ln x_i$ fall into the rejection region from part (b)?

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

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

Since $0.059 > 0.05$, **Do NOT Reject H_0 at $\alpha = 0.05$**.

ii) Part (b): Reject H_0 if $\sum_{i=1}^{5} \ln x_i \le 3.875$.

Observed $\sum_{i=1}^{5} \ln x_i \approx 4$ does NOT fall into the rejection region.

Do NOT Reject H_0 at $\alpha = 0.05$.

They agree!!!

g) Find the 90% confidence lower bound for β.

That is, find the 90% confidence interval for β of the form (a, ∞).

Hint: Recall Examples for 10/30/2020 (3):

A $(1-\alpha)\,100\,\%$ confidence interval for β: $\left(\dfrac{\chi^2_{1-\alpha/2}(4n)}{2\sum\limits_{i=1}^{n} \ln x_i}, \dfrac{\chi^2_{\alpha/2}(4n)}{2\sum\limits_{i=1}^{n} \ln x_i} \right)$.

$\left(\dfrac{\chi^2_{1-\alpha}(4n)}{2\sum\limits_{i=1}^{n} \ln x_i}, \infty \right)$ would also have a $(1-\alpha)\,100\,\%$ confidence level.

$$\chi^2_{0.90}(20) = 12.44.$$

$$\left(\dfrac{\chi^2_{1-\alpha}(4n)}{2\sum\limits_{i=1}^{n} \ln x_i}, \infty \right) = \left(\frac{12.44}{2 \cdot 4}, \infty \right) \approx \mathbf{(1.555, \infty)}.$$

h) (i) Using the p-value obtained in part (e), state your decision (Reject H_0 or Do NOT Reject H_0) at $\alpha = 0.10$.

(ii) Does your decision agree with part (g)? That is, does the interval in part (g) cover the value $\beta = 1.4$?

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

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

Since $0.059 < 0.10$, **Reject $\mathbf{H_0}$ at $\boldsymbol{\alpha} = \mathbf{0.10}$**.

ii) $\beta = 1.4$ is NOT covered by the confidence interval in part (g). That is, based on our data set, the lowest "believable" (with 90% confidence) value of β is 1.555. $\beta = 1.4$ is NOT a "believable" (with 90% confidence) value of β.

Reject $\mathbf{H_0: \beta = 1.4}$ at $\boldsymbol{\alpha = 0.10}$.

They agree!!!

Bonus. Let $\theta > 0$ and let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the distribution with probability density function

$$f(x;\theta) = \frac{1}{\theta} \cdot x^{\frac{1}{\theta} - 1}, \qquad 0 < x < 1.$$

We wish to test $H_0: \theta = 4$ vs. $H_1: \theta < 4$.

Suppose $n = 3$, and $x_1 = 0.2$, $x_2 = 0.3$, $x_3 = 0.5$. Find the p-value.

Recall: Examples for 10/16/2020 (1):

$W = -\ln X$ has $\text{Exponential}(\theta) = \text{Gamma}(\alpha = 1, \theta)$ distribution.

Examples for 11/02/2020 (1):

$\prod_{i=1}^{n} X_i$ is a sufficient statistic for θ,

$\sum_{i=1}^{n} \ln X_i$ is a sufficient statistic for θ,

$-\sum_{i=1}^{n} \ln X_i$ is a sufficient statistic for θ.

$-\sum_{i=1}^{3} \ln X_i = \sum_{i=1}^{3} W_i$ has a $\text{Gamma}(\alpha = 3, \theta)$ distribution.

θ is "θ". $\quad$ $E(W) = \theta$.

Small $\theta \quad \Rightarrow \quad$ small $w = -\ln x$.

The sign is the same as the sign in H_1.

Reject H_0 if $-\sum_{i=1}^{n} \ln x_i \le c$.

OR

Let $\theta < 4$.

$$\frac{L(4)}{L(\theta)} = \left(\frac{\theta}{4}\right)^n \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{4}-\frac{1}{\theta}} \leq k.$$

$$\Leftrightarrow \qquad \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{4}-\frac{1}{\theta}} \leq k_1$$

$$\frac{1}{4}-\frac{1}{\theta} < 0 \quad \text{since} \quad \theta < 4$$

$$\Leftrightarrow \qquad -\sum_{i=1}^{n} \ln x_i \leq c.$$

Reject H_0 if $-\sum_{i=1}^{n} \ln x_i \leq c.$

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

Observed $-\sum_{i=1}^{n} \ln x_i = -\ln 0.2 - \ln 0.3 - \ln 0.5 \approx 3.5.$

P-value $= P\left(-\sum_{i=1}^{3} \ln X_i \leq 3.5 \,\middle|\, \theta = 4\right) = P(\text{Gamma}(\alpha = 3,\ \theta = 4) \leq 3.5).$

```
> pgamma(3.5,3,1/4)
[1] 0.05880372
```

OR

$$\int_0^{3.5} \frac{1}{2 \cdot 4^3} x^{3-1} e^{-\frac{x}{4}} dx = 0.05880\ldots$$

OR

$$\text{P-value} = P\left(-\sum_{i=1}^{3} \ln X_i \le 3.5 \,\middle|\, \theta = 4\right) = P(\text{Gamma}(\alpha = 3,\ \theta = 4) \le 3.5)$$

$$= P\left(\text{Poisson}\left(\frac{3.5}{4}\right) \ge 3\right) = 1 - P(\text{Poisson}(0.875) \le 2)$$

$$= 1 - \left(\frac{0.875^{0}\, e^{-0.875}}{0\,!} + \frac{0.875^{1}\, e^{-0.875}}{1\,!} + \frac{0.875^{2}\, e^{-0.875}}{2\,!}\right)$$

$$\approx \mathbf{0.0588}.$$

```
> 1-ppois(2,3.5/4)
[1] 0.05880372
```

OR

$$\text{P-value} = P\left(-\sum_{i=1}^{3} \ln X_i \le 3.5 \,\middle|\, \theta = 4\right) = P(\text{Gamma}(\alpha = 3,\ \theta = 4) \le 3.5)$$

$$= P\left(\chi^2(2 \cdot 3) \le \frac{2}{4} \cdot 3.5\right) = P(\chi^2(6) \le 1.75).$$

```
> pchisq(2*3.5/4,2*3)
[1] 0.05880372
```