STAT 410
Fall 2020

Final Exam

Friday, December 18, 8:00 – 11:00 am CST

There are 9 problems on the exam, most of them have multiple parts. The point value of each question is shown in parentheses before the question. The total number of points for the exam is 150.

Make sure that you include everything you wish to submit, and that the submission process has completed. You do not need to include the question statements with your work. However, please label you work clearly. Neatness and organization are appreciated. Please put your final answers at the end of your work and mark them clearly.

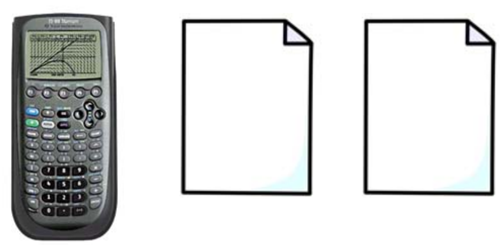
If the answer is a function, specify what kind of function it is (p.d.f., p.m.f., c.d.f., or m.g.f.), and **its support must be included**.

Be sure to show all your work; your partial credit might depend on it.

No credit will be given without supporting work.

The exam is closed book and closed notes.

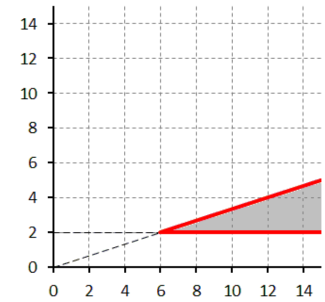
You are allowed to use a calculator and
two 8½" x 11" sheets (both sides) with notes.



You are allowed to use
https://www.wolframalpha.com/calculators/integral-calculator/
https://www.symbolab.com/solver/definite-integral-calculator
https://www.integral-calculator.com/
https://www.desmos.com/calculator

You are allowed to use
R, RStudio, and Microsoft Excel.

1. Let X and Y have the joint probability density function



$$f(x,y) = \frac{162\,y}{x^4}, \qquad y > 2, \quad x > 3\,y,$$

zero otherwise.

a) (9) Find $P(Y > 4 \mid X = 18)$.

$$f_X(x) = \int_2^{x/3} \frac{162\,y}{x^4}\,dy = \frac{81}{x^4}\left(\frac{x^2}{9} - 4\right) = \frac{9}{x^2} - \frac{324}{x^4}, \qquad x > 6.$$

$$f_{Y|X}(y\,|\,x) = \frac{\dfrac{162\,y}{x^4}}{\dfrac{81}{x^4}\left(\dfrac{x^2}{9} - 4\right)} = \frac{2\,y}{\dfrac{x^2}{9} - 4}, \qquad 2 < y < \frac{x}{3}.$$

$$f_{Y|X}(y\,|\,18) = \frac{2\,y}{32}, \qquad 2 < y < 6.$$

$$P(Y > 4 \mid X = 18) = \int_4^6 \frac{2\,y}{32}\,dy = \frac{36 - 16}{32} = \mathbf{\frac{5}{8}} = \mathbf{0.625}.$$

b) (9) Find $E(X \mid Y = y)$.

$$f_Y(y) = \int_{3y}^{\infty} \frac{162\,y}{x^4}\,dx = \frac{162\,y}{3\,(3\,y)^3} = \frac{2}{y^2}, \qquad y > 2.$$

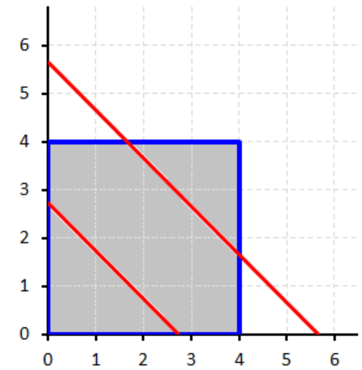
$$f_{X|Y}(x\,|\,y) = \frac{\dfrac{162\,y}{x^4}}{\dfrac{2}{y^2}} = \frac{81\,y^3}{x^4}, \qquad x > 3\,y.$$

$$E(X \mid Y = y) = \int_{3y}^{\infty} x \cdot \frac{81\,y^3}{x^4}\,dx = \left(-\frac{81\,y^3}{2\,x^2}\right)\Bigg|_{x=3y}^{x=\infty} = \frac{9\,y}{2}, \qquad y > 2.$$

2. Suppose that X has a (continuous) Uniform distribution on interval $(0, 4)$,

the p.d.f. of Y is $f_Y(y) = \dfrac{y}{8}, \quad 0 < y < 4,$ zero otherwise,

and X and Y are independent.



a) (16) Find the probability distribution of $W = X + Y$.

$$f_W(w) = f_{X+Y}(w) = \int_{-\infty}^{\infty} f_X(w-y) \cdot f_Y(y)\, dy$$

$$f_X(x) = \frac{1}{4}, \qquad 0 < x < 4. \qquad\qquad f_Y(y) = \frac{y}{8}, \qquad 0 < y < 4.$$

$$f_X(w-y) = \frac{1}{4}, \qquad 0 < w-y < 4.$$

$$w - 4 < y < w.$$

$$y > 0 \quad \& \quad y > w-4 \qquad\qquad y < 4 \quad \& \quad y < w.$$

Case 1: $\quad 0 < w < 4. \qquad$ Then $\quad w - 4 < 0 \quad \& \quad w < 4.$

$$f_W(w) = f_{X+Y}(w) = \int_0^w \frac{1}{4} \cdot \frac{y}{8}\, dy = \frac{1}{64} w^2, \qquad 0 < w < 4.$$

Case 2: $\quad 4 < w < 8. \qquad$ Then $\quad 0 < w - 4 \quad \& \quad 4 < w.$

$$f_W(w) = f_{X+Y}(w) = \int_{w-4}^{4} \frac{1}{4} \cdot \frac{y}{8}\, dy = \frac{1}{64}\left[16 - (w-4)^2 \right] = \frac{8w - w^2}{64},$$

$$4 < w < 8.$$

OR

$$f_W(w) = f_{X+Y}(w) = \int_{-\infty}^{\infty} f_X(x) \cdot f_Y(w-x)\, dx$$

Case 1: $0 < w < 4$.

$$f_W(w) = f_{X+Y}(w) = \int_0^w \frac{1}{4} \cdot \frac{w-x}{8}\, dx = \ldots, \qquad 0 < w < 4.$$

Case 2: $4 < w < 8$.

$$f_W(w) = f_{X+Y}(w) = \int_{w-4}^4 \frac{1}{4} \cdot \frac{w-x}{8}\, dx = \ldots, \qquad 4 < w < 8.$$

OR

Case 1: $0 < w < 4$.

$$F_W(w) = P(W \le w) = P(X+Y \le w) = \int_0^w \left(\int_0^{w-y} \frac{1}{4} \cdot \frac{y}{8}\, dx \right) dy$$

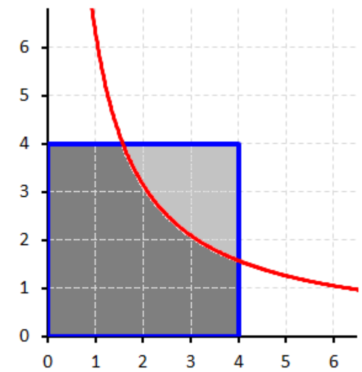
$$= \int_0^w \frac{wy - y^2}{32}\, dy = \frac{w^3}{64} - \frac{w^3}{96} = \frac{w^3}{192}, \qquad 0 < w < 4.$$

Case 2: $4 < w < 8$.

$$F_W(w) = P(W \le w) = P(X+Y \le w) = 1 - \int_{w-4}^4 \left(\int_{w-y}^4 \frac{1}{4} \cdot \frac{y}{8}\, dx \right) dy$$

$$= 1 - \int_{w-4}^4 \frac{4y - wy + y^2}{32}\, dy = \frac{-w^3 + 12w^2 - 64}{192}, \qquad 4 < w < 8.$$

b) (9) What is the probability distribution of $V = X \cdot Y$.



$$F_V(v) = P(X \cdot Y \le v)$$

$$= 1 - \int_{v/4}^{4} \left(\int_{v/y}^{4} \frac{1}{4} \cdot \frac{y}{8}\, dx \right) dy$$

$$= 1 - \int_{v/4}^{4} \left(\frac{y}{8} - \frac{v}{32} \right) dy$$

$$= 1 - \left. \frac{y^2}{16} \right|_{v/4}^{4} + \left. \frac{v \cdot y}{32} \right|_{v/4}^{4} = \frac{v^2}{256} + \frac{v}{8} - \frac{v^2}{128}$$

$$= \frac{v}{8} - \frac{v^2}{256}, \qquad 0 < v < 16.$$

OR

$X = X, \quad V = X \cdot Y.$ Then $X = X, \quad Y = \dfrac{V}{X}$.

$0 < x < 4 \quad \Rightarrow \quad 0 < x < 4.$

$0 < y < 4 \quad \Rightarrow \quad 0 < \dfrac{v}{x} < 4 \quad \Rightarrow \quad 0 < v < 4x \quad \& \quad x > \dfrac{v}{4}.$

$$J = \begin{vmatrix} 1 & 0 \\ -\dfrac{v}{x^2} & \dfrac{1}{x} \end{vmatrix} = \frac{1}{x}.$$

$$f_{X,V}(x,v) = f_{X,Y}(x, \frac{v}{x}) \cdot |J| = \left(\frac{1}{4} \cdot \frac{v}{8x} \right) \cdot \frac{1}{x} = \frac{v}{32x^2}.$$

$$f_V(v) = \int_{-\infty}^{\infty} f_{X,V}(x,v)\,dx = \int_{v/4}^{4} \frac{v}{32x^2}\,dx = -\frac{v}{32x}\Bigg|_{v/4}^{4}$$

$$= \frac{1}{8} - \frac{v}{128}, \qquad 0 < v < 16.$$

OR

$Y = Y, \quad V = X \cdot Y.$ Then $X = \frac{V}{Y}, \quad Y = Y.$

$$0 < x < 4 \quad \Rightarrow \quad 0 < \frac{v}{y} < 4 \quad \Rightarrow \quad 0 < v < 4y \quad \& \quad y > \frac{v}{4}.$$

$$0 < y < 4 \quad \Rightarrow \quad 0 < y < 4.$$

$$J = \begin{vmatrix} \frac{1}{y} & -\frac{v}{y^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{y}.$$

$$f_{V,Y}(v,y) = f_{X,Y}(\frac{v}{y}, y) \cdot |J| = \left(\frac{1}{4} \cdot \frac{y}{8} \right) \cdot \frac{1}{y} = \frac{1}{32}.$$

$$f_V(v) = \int_{-\infty}^{\infty} f_{V,Y}(v,y)\,dy = \int_{v/4}^{4} \frac{1}{32}\,dy = \frac{y}{32}\Bigg|_{v/4}^{4}$$

$$= \frac{1}{8} - \frac{v}{128}, \qquad 0 < v < 16.$$

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

$$f_X(x) = \frac{1}{2\sqrt{\beta x}}, \qquad 0 < x < \beta, \qquad \text{zero elsewhere.}$$

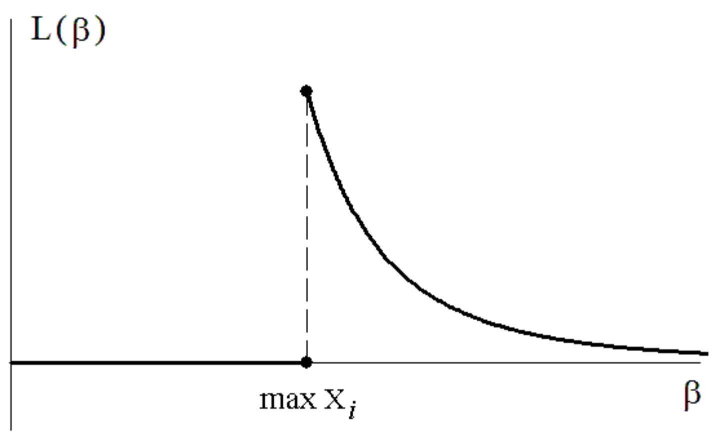
a) (9) (i) Find the maximum likelihood estimator for β, $\hat{\beta}$.

(ii) Suppose $n = 5$, and $x_1 = 0.09$, $x_2 = 0.36$, $x_3 = 1.96$, $x_4 = 6.25$, $x_5 = 10.24$. Find the maximum likelihood estimate for β, $\hat{\beta}$.

$$L(\beta) = \prod_{i=1}^{n} \left(\frac{1}{2\sqrt{\beta x_i}} \right) = \frac{1}{2^n \beta^{n/2}} \cdot \prod_{i=1}^{n} \frac{1}{\sqrt{x_i}}, \qquad \beta > \max x_i,$$

$$L(\beta) = 0, \qquad \beta < \max x_i.$$

$$\ln L(\beta) = -n \ln 2 - \frac{n}{2} \ln \beta - \frac{1}{2} \sum_{i=1}^{n} \ln x_i. \qquad \frac{d}{d\beta} \ln L(\beta) = -\frac{n}{2\beta} = 0 \quad ???$$



The goal is to find where the maximum of $L(\beta)$ occurs.

Therefore,

$$\hat{\beta} = \mathbf{\max X_i}.$$

$x_1 = 0.09$, $x_2 = 0.36$, $x_3 = 1.96$, $x_4 = 6.25$, $x_5 = 10.24$.

$\Rightarrow \quad \hat{\beta} = \max x_i = \mathbf{10.24}$.

b) (7) Is the maximum likelihood estimator $\hat{\beta}$ an unbiased estimator of β?

If $\hat{\beta}$ is not an unbiased estimator of β, construct an unbiased estimator of β from $\hat{\beta}$.

$$F_X(x) = P(X \le x) = \int_0^x \frac{1}{2\sqrt{\beta u}}\,du = \left(\frac{\sqrt{u}}{\sqrt{\beta}}\right)\Bigg|_0^x = \frac{\sqrt{x}}{\sqrt{\beta}} = \left(\frac{x}{\beta}\right)^{1/2},$$

$$0 \le x < \beta.$$

$$F_{\max X_i}(x) = [F_X(x)]^n = \left(\frac{\sqrt{x}}{\sqrt{\beta}}\right)^n = \left(\frac{x}{\beta}\right)^{n/2}, \qquad 0 \le x < \beta.$$

$$f_{\max X_i}(x) = F'_{\max X_i}(x) = \frac{n\,x^{(n-2)/2}}{2\,\beta^{n/2}}, \qquad 0 < x < 4.$$

$$E(\max X_i) = \int_0^\beta x \cdot \frac{n\,x^{(n-2)/2}}{2\,\beta^{n/2}}\,dx = \frac{n}{2\,\beta^{n/2}} \cdot \int_0^\beta x^{n/2}\,dx$$

$$= \frac{n}{2\,\beta^{n/2}} \cdot \frac{2\,x^{(n+2)/2}}{(n+2)}\Bigg|_0^\beta = \frac{n}{n+2}\,\beta \neq \beta.$$

$\hat{\beta} = \max X_i$ is NOT an unbiased estimator of β.

Consider $\quad \hat{\hat{\beta}} = \dfrac{n+2}{n}\,\hat{\beta} = \dfrac{n+2}{n}\max X_i.$

$E(\hat{\hat{\beta}}) = \beta. \qquad \hat{\hat{\beta}} = \dfrac{n+2}{n}\max X_i$ is an unbiased estimator of β.

4. (7) Suppose that the crude oil prices per barrel (X) and the price of regular unleaded gasoline in Anytown (Y) follow a bivariate normal distribution with

$$\mu_X = \$45, \quad \sigma_X = \$2, \quad \mu_Y = \$2.10, \quad \sigma_Y = \$0.10, \quad \rho = 0.28.$$

What is the probability that the price of regular unleaded gasoline in Anytown is below \$2.07, if the crude oil price is \$48 per barrel? That is, find $P(Y < 2.07 \mid X = 48)$.

Given $X = 48$, Y has Normal distribution

with mean $\mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) = 2.10 + 0.28 \cdot \frac{0.10}{2} \cdot (48 - 45) = 2.142$

and variance $\left(1 - \rho^2\right) \cdot \sigma_Y^2 = \left(1 - 0.28^2\right) \cdot 0.10^2 = 0.009216$

(standard deviation = 0.096).

$$P(Y < 2.07 \mid X = 48) = P\left(Z < \frac{2.07 - 2.142}{0.096}\right) = P(Z < -0.75) = \mathbf{0.2266}.$$

5. Let $\delta > 0$ and let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution with probability density function

$$f(x;\delta) = \frac{\delta}{(1+x)^{\delta+1}}, \qquad x > 0, \qquad \text{zero otherwise.}$$

a) (7) Show that $W = \ln(1+X)$ follows a Gamma distribution.

What are the parameters α and θ for this Gamma distribution?

No credit will be given without proper justification.

$$W = \ln(1+X) \qquad x = e^{w} - 1 \qquad \frac{dx}{dw} = e^{w}$$

$$f_W(w) = \delta\, e^{-(\delta+1)w} \cdot \left| e^{w} \right| = \delta\, e^{-\delta w}, \qquad w > 0.$$

$\Rightarrow$ W has Exponential distribution with mean $\frac{1}{\delta}$.

Exponential with mean $\frac{1}{\delta}$ = Gamma$(\alpha = 1,\ \theta = \frac{1}{\delta})$. $(\lambda = \delta)$

OR

$$F_X(x) = \int_0^x \frac{\delta}{(1+u)^{\delta+1}}\,du = -\frac{1}{(1+u)^{\delta}}\Bigg|_0^x = 1 - \frac{1}{(1+x)^{\delta}}, \qquad x > 0.$$

$$F_W(w) = P(\ln(1+X) \le w) = P(X \le e^{w} - 1) = F_X(e^{w} - 1) = 1 - e^{-\delta w},$$

$$w > 0.$$

$\Rightarrow$ W has Exponential distribution with mean $\frac{1}{\delta}$.

Exponential with mean $\frac{1}{\delta}$ = Gamma$(\alpha = 1,\ \theta = \frac{1}{\delta})$. $(\lambda = \delta)$

b) (2) Find a sufficient statistic $Y = u(X_1, X_2, \ldots, X_n)$ for δ.

$$f(x_1;\delta)\, f(x_2;\delta)\, \ldots\, f(x_n;\delta) = \prod_{i=1}^{n} \left(\frac{\delta}{(1+x_i)^{\delta+1}} \right) = \frac{\delta^n}{\left(\prod_{i=1}^{n} (1+x_i) \right)^{\delta+1}}$$

$\Rightarrow$ By Factorization Theorem, $\prod_{i=1}^{n} (1+X_i)$ is a sufficient statistic for δ.

OR

$f(x;\delta) = \exp\left\{ -(\delta+1)\cdot\ln(1+x) + \ln\delta \right\}$. $\Rightarrow$ $K(x) = \ln(1+x)$.

$\Rightarrow$ $\sum_{i=1}^{n} K(X_i) = \sum_{i=1}^{n} \ln(1+X_i)$ is a sufficient statistic for δ.

c) (9) (i) Obtain the maximum likelihood estimator for δ, $\hat{\delta}$.

(ii) Suppose $n = 4$, and $x_1 = 0.03$, $x_2 = 0.3$, $x_3 = 2.0$, $x_4 = 4.0$.

Obtain the maximum likelihood estimate for δ, $\hat{\delta}$.

$$L(\delta) = \prod_{i=1}^{n} \left(\frac{\delta}{(1+x_i)^{\delta+1}} \right) = \frac{\delta^n}{\left(\prod_{i=1}^{n} (1+x_i) \right)^{\delta+1}}.$$

$$\ln L(\delta) = n \cdot \ln \delta - (\delta+1) \cdot \sum_{i=1}^{n} \ln(1+x_i).$$

$$(\ln L(\delta))' = \frac{n}{\delta} - \sum_{i=1}^{n} \ln(1+x_i) = 0.$$

$$\Rightarrow \quad \hat{\delta} = \frac{n}{\sum_{i=1}^{n} \ln(1+X_i)}.$$

$n = 4$ $\qquad \ln 1.03 + \ln 1.3 + \ln 3 + \ln 5 \approx 3.$

$$\hat{\delta} = \frac{\mathbf{4}}{\mathbf{3}} \approx 1.3333.$$

d) (7) Is the maximum likelihood estimator $\hat{\delta}$ an unbiased estimator of δ?

If $\hat{\delta}$ is not an unbiased estimator of δ, construct an unbiased estimator of δ from $\hat{\delta}$.

$Y = \sum_{i=1}^{n} W_i = \sum_{i=1}^{n} \ln\left(1+X_i\right)$ has a Gamma$\left(\alpha = n, \theta = \frac{1}{\delta}\right)$ distribution.

$$\hat{\delta} = \frac{n}{\sum_{i=1}^{n} \ln(1+X_i)} = \frac{n}{Y}.$$

If T_α has a Gamma$(\alpha, \theta = 1/\lambda)$ distribution, then

$$E(T_\alpha^m) = \frac{\theta^m \, \Gamma(\alpha+m)}{\Gamma(\alpha)} = \frac{\Gamma(\alpha+m)}{\lambda^m \, \Gamma(\alpha)}, \qquad m > -\alpha.$$

$$E\left(\frac{1}{Y}\right) = E\left(Y^{-1}\right) = \frac{\Gamma(n-1)}{\delta^{-1}\,\Gamma(n)} = \frac{\delta}{n-1}.$$

$$E(\hat{\delta}) = E\left(\frac{n}{Y}\right) = n\,E\left(\frac{1}{Y}\right) = n \cdot \frac{\delta}{n-1} = \frac{n}{n-1} \cdot \delta = \delta + \frac{\delta}{n-1} \neq \delta.$$

$\hat{\delta}$ is NOT an unbiased estimator of δ. $\quad$ $\text{bias}(\hat{\delta}) = E(\hat{\delta}) - \delta = \dfrac{\delta}{n-1}$.

Consider $\hat{\hat{\delta}} = \dfrac{n-1}{n} \cdot \hat{\delta} = \dfrac{n-1}{\sum_{i=1}^{n} \ln(1+X_i)}$.

Then $E(\hat{\hat{\delta}}) = \dfrac{n-1}{n} \cdot E(\hat{\delta}) = \delta$. $\quad$ $\hat{\hat{\delta}}$ is an unbiased estimator of δ.

e) (9) Suppose $n = 4$, and $x_1 = 0.03$, $x_2 = 0.3$, $x_3 = 2.0$, $x_4 = 4.0$.

Construct a 90% confidence interval for δ.

"Hint": Use $\sum_{i=1}^{n} \ln\left(1+X_i\right)$.

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

$$\Rightarrow \quad 2\,\delta \sum_{i=1}^{n} \ln(1+X_i) \text{ has a } \chi^2(2\alpha = 2n) \text{ distribution.}$$

$$\Rightarrow \quad P\left(\chi^2_{1-\alpha/2}(2n) < 2\,\delta \sum_{i=1}^{n} \ln(1+X_i) < \chi^2_{\alpha/2}(2n)\right) = 1-\alpha.$$

$$\Rightarrow \quad P\left(\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n} \ln(1+X_i)} < \delta < \frac{\chi^2_{\alpha/2}(2n)}{2\sum_{i=1}^{n} \ln(1+X_i)}\right) = 1-\alpha.$$

A $(1-\alpha)\,100\,\%$ confidence interval for δ

$$\left(\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n} \ln(1+X_i)}, \frac{\chi^2_{\alpha/2}(2n)}{2\sum_{i=1}^{n} \ln(1+X_i)}\right).$$

$n = 4$ $\qquad$ $\ln 1.03 + \ln 1.3 + \ln 3 + \ln 5 \approx 3.$

$\chi^2_{0.95}(8) = 2.733,$ $\qquad$ $\chi^2_{0.05}(8) = 15.51.$

$\left(\frac{2.733}{2\cdot 3}, \frac{15.51}{2\cdot 3}\right)$ $\qquad$ $\mathbf{(0.4555, 2.585)}$

6. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with probability density function

$$f(x \mid \delta) = \frac{\delta}{(1+x)^{\delta+1}}, \qquad x > 0, \qquad \delta > 0, \qquad \text{zero otherwise.}$$

Let the prior p.d.f. of δ be $\text{Gamma}(\alpha, \theta)$. That is,

$$\pi(\delta) = \frac{1}{\Gamma(\alpha)\,\theta^{\alpha}}\,\delta^{\alpha-1}\,e^{-\delta/\theta}, \qquad \delta > 0.$$

a) (7) Find the posterior distribution of δ, given $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$. What are the parameters of this distribution?

$$f(x_1, x_2, \ldots x_n \mid \delta) = \prod_{i=1}^{n} \frac{\delta}{(1+x_i)^{\delta+1}} = \delta^n \left(\prod_{i=1}^{n} (1+x_i) \right)^{-\delta-1}.$$

$$= \left(\prod_{i=1}^{n} (1+x_i) \right)^{-1} \delta^n\, e^{-\delta \sum_{i=1}^{n} \ln(1+x_i)}$$

$$f(x_1, x_2, \ldots x_n, \delta) = f(x_1, x_2, \ldots x_n \mid \delta) \times \pi(\delta)$$

$$= \left(\prod_{i=1}^{n} (1+x_i) \right)^{-1} \delta^n\, e^{-\delta \sum_{i=1}^{n} \ln(1+x_i)} \times \frac{1}{\Gamma(\alpha)\,\theta^{\alpha}}\,\delta^{\alpha-1}\,e^{-\delta/\theta}$$

$$= \ldots\ \delta^{n+\alpha-1}\, e^{-\delta\left(\sum_{i=1}^{n} \ln(1+x_i) + \frac{1}{\theta} \right)}.$$

$\Rightarrow$ the posterior distribution of δ, given $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$,

is **Gamma** with New $\alpha = n + \alpha$ and New $\theta = \dfrac{1}{\sum_{i=1}^{n} \ln(1+x_i) + \frac{1}{\theta}}$.

b) (6) Show that $\mathrm{E}(\delta \mid \mathrm{X}_1 = x_1, \mathrm{X}_2 = x_2, \ldots, \mathrm{X}_n = x_n)$ is a weighted average of the maximum likelihood estimator $\hat{\delta}$ and the prior mean $\alpha\,\theta$. (What are the weights?)

$\Rightarrow$ (conditional mean of δ, given $\mathrm{X}_1 = x_1, \mathrm{X}_2 = x_2, \ldots, \mathrm{X}_n = x_n$)

$$= (\text{New } \alpha) \times (\text{New } \theta) = \frac{n+\alpha}{\sum\limits_{i=1}^{n} \ln(1+x_i) + \frac{1}{\theta}}.$$

$$\frac{n+\alpha}{\sum\limits_{i=1}^{n} \ln(1+x_i) + \frac{1}{\theta}}$$

$$= \frac{n}{\sum\limits_{i=1}^{n} \ln(1+x_i)} \cdot \frac{\sum\limits_{i=1}^{n} \ln(1+x_i)}{\sum\limits_{i=1}^{n} \ln(1+x_i) + \frac{1}{\theta}} + \alpha\,\theta \cdot \frac{\frac{1}{\theta}}{\sum\limits_{i=1}^{n} \ln(1+x_i) + \frac{1}{\theta}}.$$

$$= \hat{\delta} \cdot \frac{\sum\limits_{i=1}^{n} \ln(1+x_i)}{\sum\limits_{i=1}^{n} \ln(1+x_i) + \frac{1}{\theta}} + (\alpha\,\theta) \cdot \frac{\frac{1}{\theta}}{\sum\limits_{i=1}^{n} \ln(1+x_i) + \frac{1}{\theta}}.$$

For fun:

$$\left(\frac{\chi^2_{1-\gamma/2}(2n+2\alpha)}{2\sum\limits_{i=1}^{n} \ln(1+x_i) + \frac{2}{\theta}},\ \frac{\chi^2_{\gamma/2}(2n+2\alpha)}{2\sum\limits_{i=1}^{n} \ln(1+x_i) + \frac{2}{\theta}} \right)$$

is a $(1-\gamma)\,100\%$ credible interval for δ, given $\mathrm{X}_1 = x_1, \mathrm{X}_2 = x_2, \ldots, \mathrm{X}_n = x_n$.

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

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

Recall (Exam 2): $\quad W = \sqrt{X}$ has Gamma$\left(\alpha = 2,\ \theta = \frac{1}{\sqrt{\beta}}\right)$ distribution,

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

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

a) (9) Suppose $n = 5$. Find the uniformly most powerful rejection region with $\alpha = 0.10$.

Let $\beta < 4$.

$$\frac{L(4;\, x_1, x_2, \ldots, x_n)}{L(\beta;\, x_1, x_2, \ldots, x_n)} = \frac{\prod_{i=1}^{n} f(x_i;4)}{\prod_{i=1}^{n} f(x_i;\beta)} = \frac{\prod_{i=1}^{n}\left(\frac{4}{2}\, e^{-\sqrt{4x_i}}\right)}{\prod_{i=1}^{n}\left(\frac{\beta}{2}\, e^{-\sqrt{\beta x_i}}\right)}$$

$$= \left(\frac{4}{\beta}\right)^n e^{\left(\sqrt{\beta}-2\right)\sum_{i=1}^{n}\sqrt{x_i}} \le k.$$

Since $\beta < 4$,

$$\left(\frac{4}{\beta}\right)^n e^{\left(\sqrt{\beta}-2\right)\sum_{i=1}^{n}\sqrt{x_i}} \le k \qquad \Leftrightarrow \qquad \sum_{i=1}^{n}\sqrt{x_i} \ge c.$$

Intuition: $\quad \sqrt{\beta}$ is "λ". $\quad$ Small $\beta \Rightarrow$ large $\sqrt{X}$.

The sign is opposite from the sign in H_1.

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

$$\Rightarrow \qquad 2\sqrt{\beta} \sum_{i=1}^{n} \sqrt{X_i} \text{ has a } \chi^2(2\alpha = 4n) \text{ distribution.}$$

$$\Rightarrow \qquad 2\sqrt{\beta} \sum_{i=1}^{5} \sqrt{X_i} \text{ has a } \chi^2(20) \text{ distribution.}$$

$$0.10 = \alpha = P\left(\sum_{i=1}^{5} \sqrt{X_i} \ge c \mid \beta = 4\right)$$

$$= P\left(2\sqrt{\beta} \sum_{i=1}^{3} \sqrt{X_i} \ge 2\sqrt{\beta}\, c \mid \beta = 4\right) = P\left(\chi^2(20) \ge 4c\right).$$

$$\Rightarrow \qquad 4c = \chi^2_{0.10}(20) = 28.41. \qquad \Rightarrow \qquad c = \mathbf{7.1025}.$$

```
> qchisq(0.90,20)
[1] 28.41198
> qchisq(0.90,20)/4
[1] 7.102995
>
> qgamma(0.90,10,2)
[1] 7.102995
```

$c \approx \mathbf{7.1}$.

The uniformly most powerful rejection region is

"Reject H_0 if $\sum_{i=1}^{3} \sqrt{x_i} \ge 7.1$."

b) (6) Suppose $n = 5$, and $x_1 = 0.09$, $x_2 = 0.36$, $x_3 = 1.96$, $x_4 = 6.25$, $x_5 = 10.24$.

Find the p-value of the test.

$$\sum_{i=1}^{n} \sqrt{x_i} = 8.$$

p-value = P(… as extreme or more extreme than the observed … | H_0 true).

$$\text{p-value} = \mathrm{P}\left(\sum_{i=1}^{5} \sqrt{\mathrm{X}_i} \ge 8 \,\middle|\, \beta = 4 \right) = \mathrm{P}\left(\mathrm{Gamma}\left(\alpha = 10,\ \theta = \frac{1}{2} \right) \ge 8 \right).$$

```
> 1-pgamma(8,10,2)
[1] 0.04329832
```

$$\text{p-value} = \mathrm{P}\left(\sum_{i=1}^{5} \sqrt{\mathrm{X}_i} \ge 8 \,\middle|\, \beta = 4 \right) = \mathrm{P}\left(\mathrm{Gamma}\left(\alpha = 10,\ \theta = \frac{1}{2} \right) \ge 8 \right)$$

$$= \mathrm{P}\left(\mathrm{Poisson}\left(2 \cdot 8 \right) \le 10 - 1 \right) = \mathrm{P}\left(\mathrm{Poisson}\left(16 \right) \le 9 \right) = \mathbf{0.043}.$$

```
> ppois(9,16)
[1] 0.04329832
```

$$\text{p-value} = \mathrm{P}\left(\sum_{i=1}^{5} \sqrt{\mathrm{X}_i} \ge 8 \,\middle|\, \beta = 4 \right) = \mathrm{P}\left(\mathrm{Gamma}\left(\alpha = 10,\ \theta = \frac{1}{2} \right) \ge 8 \right)$$

$$= \mathrm{P}\left(\chi^2(20) \ge 4 \cdot 8 \right) = \mathrm{P}\left(\chi^2(20) \ge 32 \right).$$

```
> 1-pchisq(32,20)
[1] 0.04329832
```

p-value $\approx$ **0.043**.

8. Suppose a random sample of size $n = 25$ is taken from a normal distribution with $\sigma = 8$ for the purpose of testing $H_0: \mu = 50$ vs. $H_1: \mu < 50$ at a 5% level of significance.

a) (3) What is the p-value of this test if the observed value of the sample mean is $\bar{x} = 46.8$?

Test Statistic: $$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{46.8 - 50}{8 / \sqrt{25}} = -2.00.$$

Left-tailed. P-value $= P(Z \le -2.00) =$ **0.0228**.

b) (6) What is the power of this test if $\mu = 47.8$?

Reject H_0 if $$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} < -z_\alpha. \qquad Z = \frac{\bar{X} - 50}{8 / \sqrt{25}} < -1.645.$$

$$\bar{X} < 50 - 1.645 \cdot \frac{8}{\sqrt{25}} = 47.368.$$

$$\text{Power}(\mu = 47.8) = P(\text{Reject } H_0 \mid \mu = 47.8) = P(\bar{X} < 47.368 \mid \mu = 47.8)$$

$$= P\left(Z < \frac{47.368 - 47.8}{8 / \sqrt{25}} \right) = P(Z < -0.27) = \mathbf{0.3936}.$$

9. 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) (5) Find the best rejection region with the significance level α of the test closest to 0.02.

Reject H_0 if $x_1 + x_2 + x_3 + x_4 \ge c$.

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

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

$$= P(\text{Poisson}(4 \cdot 3) \ge c) = P(\text{Poisson}(12) \ge c) \approx 0.02.$$

$P(\text{Poisson}(12) \le 19) = 0.979.$ $\quad$ $P(\text{Poisson}(12) \ge 20) = 0.021.$

Reject H_0 if $\mathbf{X_1 + X_2 + X_3 + X_4 \ge 20}$.

b) (4) Find the power of the test from part (a) if $\lambda = 4$.

$$\text{Power} = P(\text{Reject } H_0 \mid H_0 \text{ is NOT true}) = P(X_1 + X_2 + X_3 + X_4 \ge 20 \mid \lambda = 4)$$

$$= P(\text{Poisson}(4\lambda) \ge 20 \mid \lambda = 4) = 1 - P(\text{Poisson}(16) \le 19)$$

$$= 1 - 0.812 = \mathbf{0.188}.$$

c) (4) Suppose $x_1 = 5,\ x_2 = 2,\ x_3 = 3,\ x_4 = 7$. Find the p-value of the test.

$$x_1 + x_2 + x_3 + x_4 = 17.$$

$$\text{P-value} = P(\text{value of } \textstyle\sum_{i=1}^{n=4} X_i \text{ as extreme or more extreme than } 17 \mid H_0 \text{ true})$$

$$= P(X_1 + X_2 + X_3 + X_4 \ge 17 \mid \lambda = 3)$$

$$= P(\text{Poisson}(4\lambda) \ge 17 \mid \lambda = 3) = 1 - P(\text{Poisson}(12) \le 16)$$

$$= 1 - 0.899 = \mathbf{0.101}.$$