Math 37600 PR - (19364) - Homework Post

Ethan Akin
Email: eakin@ccny.cuny.edu

Fall, 2024

Contents

Sec. 1.4

1.4.8: For a randomly chosen spring
$P(I) = .30, P(II) = .25, P(III) = .45$. Let D be the event that the chosen spring is defective.
$P(D|I) = .01, P(D|II) = .04, P(D|III) = .02$.

(a) $P(D) = P(D|I)P(I) + P(D|II)P(II) + P(D|III)P(III) = .003 + .01 + .009 = .022$.

(b) $P(II|D) = P(D|II)P(II) \div P(D) = .01 \div .022 = .45$.

1.4.10: $P(C_1) = \frac{2}{3}, P(C_2) = \frac{1}{3}$. $P(C|C_1) = \frac{3}{10}, P(C|C_2) = \frac{8}{10}$.

So $P(C) = P(C|C_1)P(C_1) + P(C|C_2)P(C_2) = \frac{14}{30}$.

$P(C_1|C) = P(C|C_1)P(C_1) \div P(C) = \frac{6}{14}$.
$P(C_2|C) = 1 - P(C_1|C)$.

Sec. 1.5, 1.6

1.5.5: Hypergeometric (a)$f_X(k) = \binom{13}{k}\binom{39}{5-k} \div \binom{52}{5}$ for $k = 0, 1, \ldots, 5$.

(b) $P(X \leq 1) = f_X(0) + f_X(1)$.

1.5.6:
$P(D_1) = \int_0^1 2x/9\, dx = 1/9, P(D_2) = \int_2^3 2x/9\, dx = 1-(4/9)$.

1.6.2: For $k = 1, \dots, 10$

$$f_X(k) = \frac{9_{k-1} \cdot 1}{10_k} = \frac{9 \cdot 8 \dots (9 - (k-2))}{10 \cdot 9 \dots (10 - (k-1))} = \frac{1}{10}.$$

Why is it $\frac{1}{10}$ for all k? Arrange the balls with the red ball first. $X = k$ if when the balls are rearranged with each of the 10! rearrangements equally likely, the red ball is in the k^{th} place. Instead of building the rearrangement by choosing a ball for each place, choose a place for each ball. $X = k$ if the place chosen for the red ball is the k^{th} place. There are 10 equally likely places for it and so $P(X = k) = \frac{1}{10}$.

For the ten balls originally listed $1, 2, \ldots, 10$, let $X_1, X_2, \ldots$ be the places in a rearrangement with all rearrangements equally likely. So $P(X_1 = i) = \frac{1}{10}$. It is tempting to say $P(X_2 = k) = \frac{1}{9}$, but that is wrong. What is true is:

$$P(X_2 = k | X_1 = i) = \begin{cases} 1/9 \text{ for } & k \neq i, \\ 0 & \text{for } k = i. \end{cases}$$

In fact $P(X_2 = k) = \frac{1}{10}$ because if we don't know what X_1 was the places for ball 2 are all equally likely.

Observe the Law of Total Probability:

$$P(X_2 = k) = \sum_{i=1}^{10} P(X_2 = k | X_1 = i) \cdot P(X_1 = i)$$

$$= 9 \cdot \frac{1}{9} \cdot \frac{1}{10} \quad + \quad 0 \cdot \frac{1}{10} = \frac{1}{10}.$$

with the 0 where $i = k$. We will see this again when we look at order statistics

1.6.9: $Y = 1$ when $X = \pm 1$ and $Y = 0$ when $X = 0$. So $f_Y(1) = \frac{2}{3}, f_Y(0) = \frac{1}{3}$.

Sec. 1.7

1.7.6: $|X| < 1$ if and only if $-1 < X < 1$. $X^2 < 9$ if and only if $-3 < X < 3$.
(a) $P(|X| < 1) = \int_{-1}^{1} \frac{x^2}{18} dx = 2\int_0^1 \frac{x^2}{18} dx = \frac{1}{27}$. $P(X^2 < 9) = 1$.

1.7.9 (b): If m is the median, then $\frac{1}{2} = \int_0^m \, 3x^2 dx$. So $\frac{1}{2} = m^3$. m is the cube root of $\frac{1}{2}$.

1.7.12(b): $F_X(x) = 0$ for $x \leq 1$ and $F_X(x) = \int_1^x \frac{1}{t^2} dt = 1 - \frac{1}{x}$.

1.7.14: $P(X \geq 3/4 | X \geq 1/2) = \int_{3/4}^1 2x dx \div \int_{1/2}^1 2x dx = (7/16) \div (3/4) = 7/12$.

Sec. 1.8

1.8.7[1.8.6]:
$E(A) = E(X(1-X)) = \int_0^1 3x(1-x)x^2\ dx = 3/20.$

1.8.9[1.8.8]: (a) $E(1/X) = \int_0^1 (1/x)2xdx = 2.$

(b) $F_Y(y) = 0$ for $y \leq 1$. For $y > 1$, $Y \leq y$ if and only if $X \geq (1/y)$.

$$F_Y(y) = \int_{1/y}^1 2xdx = 1 - (1/y^2). \qquad f_Y(y) = F'_Y(y) = 2/y^3.$$

(c) $E(Y) = \int_1^\infty y(2/y^3)dy = 2.$

1.8.12[1.8.11]: (a) $E(X^3) = \int_0^1 x^3(3x^2)dx = 1/2.$

(b) $Y = X^3 \leq y$ if and only if $X \leq y^{1/3}$.
$F_Y(y) = \int_0^{y^{1/3}} 3x^2dx = y$. $f_Y(y) = 1$. Uniform.

Sec. 1.9

1.9.1(c): $E(X) = \int_1^\infty x(2/x^3)dx = 2.$
$E(X^2) = \int_1^\infty x^2(2/x^3)dx$ Integral diverges. No variance.

1.9.4: The difference is the variance which is non-negative.

1.9.5: If $c = 0$ then symmetry means that $f(x)$ is an even function. So $xf(x)$ is an odd function and, if the integral converges, $\int_{-\infty}^\infty xf(x)dx = 0.$

In general, if X has pdf $f_X(x)$ then $Y = X - c$ has pdf $f_Y(y) = f_X(y + c)$. So f_Y is symmetric about 0 and $E(Y) = 0.$

Sec. 3.1

3.1.15[3.1.14]: X is $Geom(p)$, $f_X(x) = pq^x$, with $p = \frac{1}{3}$. For any positive integer k, $G_X(k) = P(X \geq k) = q^k$.

So for $x \geq k$, $P(X = x|X \geq k) = pq^{x-k}$. That is, the conditional distribution of $X - k$ assuming $X \geq k$, is the same $Geom(p)$ distribution as that of X.

3.1.27[3.1.23]: X is $Geom(p)$ and so for any positive integer k, $G_X(k) = P(X \geq k) = q^k$.

$$P(X \geq k + j|X \geq k) = G_X(k + j) \div G_X(k) = q^j = G_X(j).$$

2.1.7[2.1.6]: X and Y are independent $Exp(1)$rv's. Suppose in general the X and Y are independent continuous rv's, and $Z = X + Y$. For any z:

$$F_Z(z) = P(Z \leq z) = \int_{x=-\infty}^{x=\infty} \int_{y=-\infty}^{y=z-x} f_X(x) f_Y(y) dy dx$$

$$= \int_{x=-\infty}^{x=\infty} f_X(x) F_Y(z-x) dx.$$

Differentiating with respect to z,
$f_Z(z) = F_Z'(z) = \int_{x=-\infty}^{x=\infty} f_X(x) f_Y(z-x) dx.$

Here

$$f_X(x) \;=\; e^{-x} \;\; (x > 0), \quad f_Y(y) \;=\; e^{-y} \;\; (y > 0)$$

and $= 0$ otherwise.
So

$$f_X(x) f_Y(z - x) = e^{-z} \;\; (x, z - x > 0),$$

i.e. for $0 < x < z$.

$$f_Z(z) \;=\; \int_{x=0}^{x=z} e^{-z} \, dx \;=\; z e^{-z} (0 < z).$$

In terms of Gamma distributions, $X, Y \sim Exp(1)$ and so $X, Y \sim \Gamma(1, 1)$. Since X and Y are independent, $X + Y \sim \Gamma(2, 1)$.

2.1.8[2.1.7]: For any $z \in (0, 1)$,

$$1 - F_Z(z) = G_Z(z) = P(Z \geq z)$$
$$= \int_{x=z}^{x=1} \int_{y=(z/x)}^{1} dydx = 1 - (z - z\ln(z)).$$

Differentiating with respect to z, $f_Z(z) = -G'_Z(z) = -\ln(z)$.

$$E(Z) = \int_0^1 -z\ln(z)dz = -\frac{1}{2}z^2[\ln(z) - \frac{1}{2}]|_0^1 = \frac{1}{4}.$$

Sec. 2.3

2.3.2: (a) $f_Y(y)dy = c_2 y^4 dy$ for $0 < y < 1$. So
$1 = c_2 \int_0^1 y^4 dy = c_2/5$. So $c_2 = 5$.

$f_{X|Y}(x|y)dx = c_1 x/y^2 dx$ for $0 < x < y$. So
$1 = c_1/y^2 \int_0^y x dx = c_1/2$. So $c_1 = 2$.

(b) $f_{X,Y}(x,y)dxdy = f_{X|Y}(x|y)dx \cdot f_Y(y)dy = 10xy^2$ for
$0 < x < y < 1$.

$f_X(x)dx = [10x \int_{y=x}^{y=1} y^2 dy]dx = \frac{10}{3}x[1 - x^3]$. for $0 < x < 1$.

(c) $P(\frac{1}{4} < X < \frac{1}{2} | Y = \frac{5}{8}) = 2 \cdot (\frac{8}{5})^2 \int_{x=\frac{1}{4}}^{x=\frac{1}{2}} xdx$.

(d) $P(\frac{1}{4} < X < \frac{1}{2}) = \frac{10}{3} \int_{x=\frac{1}{4}}^{x=\frac{1}{2}} x(1 - x^3)dx$.

2.3.3: $f_{X,Y}(x,y)dxdy = 21x^2y^3dxdy$ for $0 < x < y < 1$.

$$f_X(x)dx = [21x^2\int_{y=x}^{y=1} y^3dy]dx = \frac{21}{4}x^2(1-x^4)dx.$$

$$E(X^n) = \frac{21}{4}\int_0^1 x^{n+2}[1-x^4]dx =$$
$$\frac{21}{4}[\frac{1}{3+n} - \frac{1}{7+n}] = \frac{21}{(3+n)(7+n)}.$$

So $E(X) = \frac{21}{32}$ and $Var(X) = \frac{21}{45} - (\frac{21}{32})^2 = .036$

$$f_Y(y)dy = 21[y^3\int_{x=0}^{x=y} x^2dx]dy = 7y^6dy.$$

$$E(Y^n) = 7\int_0^1 y^{6+n}dy = \frac{7}{7+n}.$$

So $E(Y) = \frac{7}{8}$ and $Var(Y) = \frac{7}{9} - (\frac{7}{8})^2 = .012$

$$f_{X|Y}(x|y)dx = (21x^2y^3) \div (7y^6)dx = 3x^2/y^3dx \quad \text{for } 0 < x < y.$$

$$E(X^n|Y=y) = \int_0^y x^n \cdot (3x^2/y^3)dx = \frac{3}{3+n}y^n.$$

So $Z = E(X|Y) = \frac{3}{4}Y$. and
$Var(X|Y) = [\frac{3}{5} - (\frac{3}{4})^2]Y^2 = \frac{3}{80}Y^2$.
$E(Z) = \frac{3}{4}E(Y) = \frac{21}{32} = E(X)$. and
$Var(Z) = (\frac{3}{4})^2 Var(Y) = .007$.

Sec. 2.4, 2.5

2.4.2[2.5.2]: $f_{X,Y}(x,y) = 2e^{-x-y}dxdy$ for $0 < x < y < \infty$.
From the shape of the region the pair is dependent.

$$f_X(x)dx = 2e^{-x}[\int_{y=x}^{y=\infty} e^{-y}dy]dx = 2e^{-2x}dx.$$

$$f_Y(y)dy = 2e^{-y}[\int_{x=0}^{x=y} e^{-x}dx]dy = 2e^{-y}[1 - e^{-y}].$$

$f_{X,Y}$ is not the product of the two marginals.

Sec. 2.2, 2.7

2.7.4: $X_1 = Y_1, X_2 = Y_2 - Y_1, X_3 = Y_3 - Y_2$. Since $0 < X_1, X_2, X_3 < \infty$, $0 < Y_1 < Y_2 < Y_3 < \infty$.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

So the determinant $J = 1$.

$$f_{Y_1 Y_2 Y_3}(y_1, y_2, y_3)\ dy_1 dy_2 dy_3 =$$
$$f_{X_1 X_2 X_3}(y_1, y_2 - y_1, y_3 - y_2)\ J\ dy_1 dy_2 dy_3 =$$
$$e^{-y_1} \cdot e^{-y_2 + y_1} \cdot e^{-y_3 + y_2}\ dy_1 dy_2 dy_3 = e^{-y_3}\ dy_1 dy_2 dy_3,$$
$$\text{for} \quad 0 < y_1 < y_2 < y_3 < \infty$$

$$f_{Y_1 Y_2}(y_1, y_2)\ dy_1 dy_2 \ = \ [\int_{y_3=y_2}^{y_3=\infty} e^{-y_3}\ dy_3]\ dy_1 dy_2 \ =$$

$$e^{-y_2}\ dy_1 dy_2 \qquad \text{for} \quad 0 < y_1 < y_2 < \infty$$

$$f_{Y_2}(y_2)\ dy_2 \ = \ [\int_{y_1=0}^{y_1=y_2} e^{-y_2}\ dy_1] dy_2 \ =$$

$$y_2 e^{-y_2}\ dy_2 \qquad \text{for} \quad 0 < y_2 < \infty$$

As it is the sum of two independent $Exp(1)$ variables, $Y_2 \sim \Gamma(2, 1)$.

Sec. 3.4

3.4.1: Use the change of variables $u = -w$.

3.4.6: With $Z = \frac{X-\mu}{\sigma}$,

$$E(|Z|) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z| e^{-z^2/2} dz = \sqrt{\frac{2}{\pi}} \int_{0}^{\infty} z e^{-z^2/2} dz.$$

With substitution $u = z^2/2$, $\int_0^\infty z e^{-z^2/2} dz = 1$.

Sec. 4.4

4.1.5: (a) $P(Y_1 = X_1) = \frac{1}{n} = \frac{(n-1)!}{n!}$.

(b) $P(Y_1 = X_n, Y_2 = X_1) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$.

(c) No mean because the series is harmonic. No second moment and hence no variance.

4.4.5: $P(Y_4 \geq 3) = 1 - P(Y_4 < 3) = 1 - F_X(3)^4$.
$X \sim Exp(1)$ and so $F_X(x) = 1 - e^{-x}$.
$P(Y_4 \geq 3) = 1 - (1 - e^{-3})^4$.

4.4.6: With $f(x) = 2x$, the median $\xi = \xi_{.5}$ is given by $\int_0^\xi 2xdx = \frac{1}{2}$. so $\xi = \frac{1}{\sqrt{2}}$.

(a) For X_1, X_2, X_3 iid with any continuous pdf,
$P(Y_1 > \xi_{.5}) = P(X > \xi_{.5})^3 = (\frac{1}{2})^3$.

(b) $f_{Y_2 Y_3}(y_2, y_3) dy_2 dy_3 = 3 \cdot 2 F_X(y_2) f_X(y_2) f_X(y_3) dy_2 dy_3$. for $0 < y_2 < y_3 < 1$.

$$f_{Y_2 Y_3}(y_2, y_3) dy_2 dy_3 = 6y_2^2 \cdot 2y_2 \cdot 2y_3 dy_2 dy_3 = 24 y_2^3 y_3 dy_2 dy_3.$$

$$f_{Y_2}(y_2) dy_2 = 3 \cdot 2 F_X(y_2) f_X(y_2)(1 - F_X(y_2)) dy_2 = 12 y_2^3 (1 - y_2^2).$$

$$f_{Y_3}(y_3) dy_3 = 3 F_X(y_3)^2 f_X(y_3) dy_3 = 6 y_3^5.$$

$$E(Y_2) = \int_0^1 12y_2^4(1 - y_2^2)dy_2 = \frac{24}{35}.$$

$$E(Y_3) = \int_0^1 6y_3^6 dy_3 = \frac{6}{7} = \frac{30}{35}.$$

$$E(Y_2^2) = \int_0^1 12y_2^5(1 - y_2^2)dy_2 = \frac{1}{2}.$$

$$E(Y_3^2) = \int_0^1 6y_3^7 dy_3 = \frac{6}{8} = \frac{3}{4}.$$

$$E(Y_2 Y_3) = \int_{y_3=0}^{y_3=1} \int_{y_2=0}^{y_2=y_3} 24y_2^4 y_3^2 dy_2 dy_3 = \frac{3}{5}.$$

$$Cov(Y_2, Y_3) = \frac{3}{5} - \frac{24 \cdot 6}{35 \cdot 7} = \frac{147 - 144}{245} = \frac{3}{245} = .012.$$

$$Var(Y_2) = \frac{1}{2} - (\frac{24}{35})^2 = .030.$$

$$Var(Y_3) = \frac{3}{4} - (\frac{6}{7})^2 = .015.$$

$$Corr(Y_2, Y_3) = .012 \div \sqrt{.030 \cdot .015} = .57.$$

4.4.27: $P(Y_1 < \xi_{.5} < Y_n) = 1 - [P(\xi_{.5} < Y_1) + P(Y_n < \xi_{.5})] = 1 - 2 \cdot 2^{-n}$.

This is $\geq .99$ when $-\ln(100) \geq \ln(2) - n\ln(2)$. So when $n \geq \ln(200)/\ln(2) = 7.64$.

Sec. 4.1

4.1.1: (b) As in Example 4.1.1, $\hat{\theta} = \bar{X} = 101.15$.

(c) The median of the data is $\frac{1}{2}(53 + 58) = 55.5$.

(d) The median $\xi_{.5}$ for an $Exp(\theta)$ is given by $\frac{1}{2} = \int_0^{\xi} \frac{1}{\theta} e^{-x/\theta} dx = 1 - e^{\xi/\theta}$. So $\xi = \theta \ln(2)$. With $\hat{\theta} = 101.15, \hat{\xi} = 70.11$.

4.1.2: (b)As in Example 4.1.3, $\hat{\theta} = \bar{X} = 201.$ and $\hat{\sigma}^2 = \frac{1}{n}(\sum_i X_i^2) - \bar{X}^2 = 293.92$.

So $\hat{\sigma} = \sqrt{293.92} = 17.14$, $\widehat{(\mu/\sigma)} = \hat{\mu}/\hat{\sigma} = 11.73$.

4.1.3: (a) $f(\theta, x) = e^{-\theta} \frac{\theta^x}{x!}$. So

$$\ell(\theta; x_1, \ldots, x_n) = -n\theta + (\sum_{i=1}^{n} x_i) \ln(\theta) - \sum_{i=1}^{n} \ln(x_i!).$$

Taking the derivative we obtain $\hat{\theta} = \bar{X}$.

Because $E_\theta(\bar{X}) = \theta$, the estimate is unbiased.

(b) $\hat{\theta} = \bar{X} = 9.5$

4.1.8: The MLE is $\hat{\theta} = \bar{X} = 2.13$. So

$$\hat{p}(0) = e^{-\hat{\theta}} = .119, \quad \hat{p}(1) = e^{-\hat{\theta}}\hat{\theta} = .253,$$
$$\hat{p}(2) = e^{-\hat{\theta}}\hat{\theta}^2/2 = .270, \quad \hat{p}(3) = e^{-\hat{\theta}}\hat{\theta}^3/3! = .191,$$
$$\hat{p}(4) = e^{-\hat{\theta}}\hat{\theta}^3/4! = .102, \quad \hat{p}(5) = e^{-\hat{\theta}}\hat{\theta}^3/5! = .043.$$

$\hat{p}(\geq 6) = .022.$

Sec. 4.2

4.2.2: $\bar{x} = 101.1, n = 20$ and
$s^2 = \frac{n}{n-1}[\frac{1}{n}\sum_i X_i^2 \; - \; \bar{X}^2] = 11121.7$ So
$s = 105.5, s/\sqrt{n} = 23.6.$

The confidence interval is $(101.1 - 23.6z_{\alpha/2}, 101.1 - 23.6z_{\alpha/2})$

With $\alpha = 5\%$, $z_{\alpha/2} = 1.96$. The interval has endpoints
101.1 ± 41.2, i.e. $(59.9, 142.3)$.

4.2.3[4.2.4]: (a) If $X \sim \Gamma(1, \theta)$ then $2X/\theta \sim \Gamma(1, 2)$ and $\frac{2}{\theta}\sum_{i=1}^{n} X_i \sim \Gamma(n, 2)$ which is $\chi^2(2n)$.

(b) Choose q_α^n and Q_α^n so that with $W \sim \chi^2(n)$, $P(W < q_\alpha^n) = P(W > Q_\alpha^n) = \alpha/2$.

The confidence interval for θ is $(2\sum_{i=1}^{n} X_i/Q_\alpha^{2n}, 2\sum_{i=1}^{n} X_i/q_\alpha^{2n})$.

(c) With $2n = 40$ degrees of freedom, $q_{.025} = 24.4, Q_{.025} = 59.3$. The confidence interval is given by

$$(40 \cdot 101.1/59.3, 40 \cdot 101.1/24.4) \;=\; (68.2, 165.7).$$

4.2.6[4.2.7]: $z_{.05} = 1.65$. We want $\sigma z_{.05}/\sqrt{n} \leq 1$.
$n = (3 \cdot 1.65)^2 = 25$.

4.2.15: $z_{.025} = 1.96$. We want $\sigma z_{.025}/\sqrt{n} \leq \sigma/4$.
$n = (4 \cdot 1.95)^2 = 61$.

4.2.17: For Poisson the sample variance equals the sample mean which is 3.4, and $n = 200$. $z_{.05} = 1.65$. The confidence interval has endpoints

$$\bar{x} \pm s z_{.05}/\sqrt{n} = 3.4 \pm .22 \quad \text{i.e.} \quad (3.18, 3.62).$$

Sec. 4.5

4.5.3: $f(x:\theta) = \theta x^{\theta-1}$; $0 < x < 1, \theta \geq 1$, $C = \{(x_1, x_2) : \frac{3}{4} \leq x_1 x_2\}$. [Recall that the MLE is $\max(1, \frac{-2}{\ln(x_1 x_2)})$]. So

$$P_\theta(C) = \int_{x_1=3/4}^{x_1=1} \int_{x_2=3/4x_1}^{x_2=1} f_\theta(x_1, x_2) dx_2 dx_1.$$

The null hypothesis H_0 is $\theta = 1$ which is uniform. So the size

$$\alpha = \int_{x_1=3/4}^{x_1=1} \int_{x_2=3/4x_1}^{x_2=1} dx_2 dx_1 = 1 - (\frac{3}{4} - \frac{3}{4}\ln(\frac{3}{4})) = .034.$$

The power is

$$\int_{x_1=3/4}^{x_1=1} \int_{x_2=3/4x_1}^{x_2=1} 4x_1 x_2 dx_2 dx_1 = [1 - ((\frac{3}{4})^2 - \frac{9}{8}\ln(\frac{3}{4}))] = .113.$$

For general θ, the power function is:

$$\int_{x_1=3/4}^{x_1=1} \int_{x_2=3/4x_1}^{x_2=1} \theta^2 x_1^{\theta-1} x_2^{\theta-1} dx_2 dx_1 = 1-((3/4)^\theta - \theta(3/4)^\theta \ln(3/4)).$$

4.5.5: Recall the that MLE is $\bar{X} = \frac{1}{2}(x_1 + x_2)$.

$$\frac{L(2; x_1, x_2)}{L(1; x_1, x_2)} = \frac{1}{4} e^{(x_1+x_2)/2}.$$

So the critical region
$C = \{(x_1, x_2) : x_1 + x_2 \leq 2\ln(2), 0 < x_1, x_2 < \infty\}$.
The independent X_1, X_2 are $\Gamma(1, \theta)$ rv's and so
$X_1 + X_2 \sim \Gamma(2, \theta)$.

$$P_\theta(C) = \int_{z=0}^{z=2\ln(2)} \frac{1}{\theta^2} z e^{-z/\theta} dz$$

$$= \int_{u=0}^{u=2\ln(2)/\theta} u e^{-u} du = 1 + (\frac{2\ln(2)}{\theta} + 1) e^{-(2\ln 2)/\theta}.$$

With $\theta = 2$ this is the size .15. With $\theta = 1$ this is the power .40.

4.5.9: If $X_1, \ldots, X_n$ is a sample from a $Poiss(\theta)$ distribution then $Y = X_1 + \cdots + X_n \sim Poiss(n\theta)$.
With $C = \{(x_1, \ldots, x_n) : \sum_{i=1}^n x_i \leq 2\}$, then

$$P_\theta(C) = e^{-n\theta}(1 + n\theta + \frac{(n\theta)^2}{2}).$$

With $n = 12$ and $\theta = \frac{1}{2}$, the size is $e^{-6}(25) = .062$.

With $n = 12$ and $\theta = \frac{1}{3}$, the power is $e^{-4}(13) = .238$.

With $n = 12$ and $\theta = \frac{1}{4}$, the power is $e^{-3}(8.5) = .423$.

4.5.11: The cdf is $F_X(\theta; x) = \frac{x}{\theta}$ for $0 < x < \theta$. With $Y = \max(X_1, \ldots, X_n)$

$$F_Y(\theta; y) \;=\; F_X(\theta; y)^n \;=\; (\frac{y}{\theta})^n.$$

For the critical region $C = \{(x_1, \ldots, x_n) : y \geq c\}$,

$$P_\theta(C) \;=\; F_Y(\theta; \theta) - F_Y(\theta; c) \;=\; 1 - (\frac{c}{\theta})^n.$$

With $\theta = 1$ and size $P_1(C) = .05$, $c = .95^{1/n}$. Then the power function is

$$P_\theta(C) \;=\; 1 - \frac{.95}{\theta^n}.$$

Sec. 4.6

4.6.8: For a $Bern(\theta)$ distribution we have a sample $X_1, \ldots, X_n$. $Y = \sum_{i=1}^{n} X_i \sim Bin(n, \theta)$ with mean $n\theta$ and variance $n\theta(1-\theta)$. So the sample mean $\bar{X} = \frac{1}{n}Y$ has mean θ and variance $\theta(1-\theta)/n$.

For n large

$$Z_\theta = \frac{Y - n\theta}{\sqrt{n\theta(1-\theta)}} = \frac{\bar{X} - \theta}{\sqrt{\theta(1-\theta)/n}}$$

has approximately a standard $\mathcal{N}(0,1)$ normal distribution.

The MLE for θ is $\bar{X}$ which is an unbiased estimate.

To test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$, we use the critical region $C = \{\bar{X} > c\} = \{Y > nc\}$.

Given size α we choose c so that $P_{\theta_0}(C) \approx \alpha$.

Using the normal approximation,we want $C = \{Z_{\theta_0} > z_\alpha\}$, or equivalently,

$$\{\bar{X} > \theta_0 + z_\alpha\sqrt{\theta_0(1-\theta_0))/n}\}, \quad \text{so that}$$
$$c = \theta_0 + z_\alpha\sqrt{\theta_0(1-\theta_0))/n}.$$

For $\theta > \theta_0$ the power is given by the probability that $Y > nc$ when $Y \sim Bin(n, \theta)$. Using the normal approximation this is the probability that the standard normal is greater than $\frac{c-\theta}{\sqrt{\theta(1-\theta)/n}}$. That is,

$$1 - \Phi[z_\alpha \frac{\sqrt{\theta_0(1-\theta_0)}}{\sqrt{\theta(1-\theta)}} - \frac{\theta - \theta_0}{\sqrt{\theta(1-\theta)/n}}].$$

At least for $\theta \leq \frac{1}{2}$ this is an increasing function of θ.

Finally, for $\bar{x}$ the realized value of $\bar{X}$, the p-value is $P_{\theta_0}(\bar{X} > \bar{x})$ or in the normal approximation

$$1 - \Phi(\frac{\bar{x} - \theta_0}{\sqrt{\theta_0(1 - \theta_0)/n}}).$$

In our case, $\theta_0 = .14$ and with a sample of $n = 590, \bar{x} = 104/590 = .176$. We use $\alpha = .01$ so that $z_\alpha = 2.33$. We obtain $c = .173$.

Since $\bar{x} > c$, we reject H_0 and accept H_1.

The p-value is $1 - \Phi(2.54) = .006$ which is less $\alpha = .01$ as expected.