Chapter 11

Stuff for Students

To be blunt, many of us are lousy teachers, and our efforts to improve are feeble. So students frequently view statistics as the worst course taken in college.
Hogg (1991)

11.1 R/Splus Statistical Software

R/Splus are statistical software packages, and *R* is the free version of *Splus*.. A very useful *R* link is (www.r-project.org/#doc).

As of January 2008, the author's personal computer has Version 2.4.1 (December 18, 2006) of *R* and Splus–2000 (see Mathsoft 1999ab).

Downloading the book's R/Splus functions *sipack.txt* into *R* or *Splus*:

Many of the homework problems use *R/Splus* functions contained in the book's website (www.math.siu.edu/olive/sipack.txt) under the file name *sipack.txt*. Suppose that you download *sipack.txt* onto a disk. Enter *R* and wait for the curser to appear. Then go to the *File* menu and drag down *Source R Code*. A window should appear. Navigate the *Look in* box until it says *3 1/2 Floppy(A:)*. In the *Files of type* box choose *All files(*.*)* and then select *sipack.txt*. The following line should appear in the main *R* window.

```
> source("A:/sipack.txt")
```

Type *ls()*. About 9 *R/Splus* functions from *sipack.txt* should appear.

Alternatively, from the website (www.math.siu.edu/olive/sipack.txt), go to the *Edit* menu and choose *Select All*, then go to the *Edit* menu and choose *Copy*. Next enter *R*, go to the *Edit* menu and choose *Paste*. These commands also enter the `sipack` functions into *R*.

When you finish your *R/Splus* session, enter the command *q()*. A window asking "*Save workspace image?*" will appear. Click on *No* if you do not want to save the programs in *R*. (If you do want to save the programs then click on *Yes*.)

If you use *Splus*, the command

```
> source("A:/sipack.txt")
```

will enter the functions into *Splus*. Creating a special workspace for the functions may be useful.

This section gives tips on using *R/Splus*, but is no replacement for books such as Becker, Chambers, and Wilks (1988), Chambers (1998), Dalgaard (2002) or Venables and Ripley (2003). Also see Mathsoft (1999ab) and use the website (http://www.google.com) to search for useful websites. For example enter the search words *R documentation*.

The command *q()* gets you out of *R* or *Splus*.

The commands *help(fn)* and *args(fn)* give information about the function fn, eg if fn = rnorm.

Making functions in R and Splus is easy.

For example, type the following commands.

```
mysquare <- function(x){
# this function squares x
r <- x^2
r }
```

The second line in the function shows how to put comments into functions.

Modifying your function is easy.

Use the fix command.

fix(mysquare)

This will open an editor such as *Notepad* and allow you to make changes.

In *Splus*, the command *Edit(mysquare)* may also be used to modify the function *mysquare*.

To save data or a function in *R*, when you exit, click on *Yes* when the "*Save worksheet image?*" window appears. When you reenter *R*, type *ls()*. This will show you what is saved. You should rarely need to save anything for the material in the first thirteen chapters of this book. In *Splus*, data and functions are automatically saved. To remove unwanted items from the worksheet, eg *x*, type *rm(x)*,
pairs(x) makes a scatterplot matrix of the columns of *x*,
hist(y) makes a histogram of *y*,
boxplot(y) makes a boxplot of *y*,
stem(y) makes a stem and leaf plot of y,
scan(), source(), and *sink()* are useful on a *Unix* workstation.
To type a simple list, use *y <− c(1,2,3.5)*.
The commands *mean(y), median(y), var(y)* are self explanatory.

The following commands are useful for a scatterplot created by the command *plot(x,y)*.
lines(x,y), lines(lowess(x,y,f=.2))
identify(x,y)
abline(out$coef), abline(0,1)

The usual arithmetic operators are $2 + 4$, $3 - 7$, $8 * 4$, $8/4$, and

```
2^{10}.
```

The ith element of vector y is $y[i]$ while the ij element of matrix x is $x[i, j]$. The second row of x is $x[2,]$ while the 4th column of x is $x[, 4]$. The transpose of x is *t(x)*.

The command *apply(x,1,fn)* will compute the row means if fn = mean. The command *apply(x,2,fn)* will compute the column variances if fn = var. The commands *cbind* and *rbind* combine column vectors or row vectors with an existing matrix or vector of the appropriate dimension.

11.2 Hints and Solutions to Selected Problems

1.10. d) See Problem 1.19 with $Y = W$ and $r = 1$.

f) Use the fact that $E(Y^r) = E[(Y^\phi)^{r/\phi}] = E(W^{r/\phi})$ where $W \sim EXP(\lambda)$. Take $r = 1$.

1.11. d) Find $E(Y^r)$ for $r = 1, 2$ using Problem 1.19 with $Y = W$.

f) For $r = 1, 2$, find $E(Y^r)$ using the the fact that $E(Y^r) = E[(Y^\phi)^{r/\phi}] = E(W^{r/\phi})$ where $W \sim EXP(\lambda)$.

1.12. a) 200

b) $0.9(10) + 0.1(200) = 29$

1.13. a) $400(1) = 400$

b) $0.9E(Z) + 0.1E(W) = 0.9(10) + 0.1(400) = 49$

1.15. a) $1\frac{A}{A+B} + 0\frac{B}{A+B} = \frac{A}{A+B}$.

b) $\frac{nA}{A+B}$.

1.16. a) $g(x_o)P(X = x_o) = g(x_o)$

b) $E(e^{tX}) = e^{tx_o}$ by a).

c) $m'(t) = x_o e^{tx_o}$, $m''(t) = x_o^2 e^{tx_o}$, $m^{(n)}(t) = x_o^n e^{tx_o}$.

1.17. $m(t) = E(e^{tX}) = e^t P(X = 1) + e^{-t} P(X = -1) = 0.5(e^t + e^{-t})$.

1.18. a) $\sum_{x=0}^n x e^{tx} f(x)$

b) $\sum_{x=0}^n x f(x) = E(X)$

c) $\sum_{x=0}^n x^2 e^{tx} f(x)$

d) $\sum_{x=0}^n x^2 f(x) = E(X^2)$

e) $\sum_{x=0}^n x^k e^{tx} f(x)$

1.19. $E(W^r) = E(e^{rX}) = m_X(r) = \exp(r\mu + r^2\sigma^2/2)$ where $m_X(t)$ is the mgf of a $N(\mu, \sigma^2)$ random variable.

1.20. a) $E(X^2) = V(X) + (E(X))^2 = \sigma^2 + \mu^2$.

b) $E(X^3) = 2\sigma^2 E(X) + \mu E(X^2) = 2\sigma^2\mu + \mu(\sigma^2 + \mu^2) = 3\sigma^2\mu + \mu^3$.

1.22. $\displaystyle \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-\frac{1}{2}y^2)dy = 1$. So $\displaystyle \int_{-\infty}^{\infty}\exp(-\frac{1}{2}y^2)dy = \sqrt{2\pi}$.

1.23. $\int_\sigma^\infty f(x|\sigma,\theta)dx = 1$, so

$$\int_\sigma^\infty \frac{1}{x^{\theta+1}}dx = \frac{1}{\theta\sigma^\theta}. \tag{11.1}$$

So
$$EX^r = \int_\sigma^\infty x^r\theta\sigma^\theta\frac{1}{x^{\theta+1}}dx = \theta\sigma^\theta\int_\sigma^\infty \frac{1}{x^{\theta-r+1}}dx = \frac{\theta\sigma^\theta}{(\theta-r)\sigma^{\theta-r}}$$
by Equation 11.1. So
$$EX^r = \frac{\theta\sigma^r}{\theta-r}$$
for $\theta > r$.

1.24.
$$EY^r = \int_0^1 y^r\frac{\Gamma(\delta+\nu)}{\Gamma(\delta)\Gamma(\nu)}y^{\delta-1}(1-y)^{\nu-1}dy =$$
$$\frac{\Gamma(\delta+\nu)}{\Gamma(\delta)\Gamma(\nu)}\frac{\Gamma(\delta+r)\Gamma(\nu)}{\Gamma(\delta+r+\nu)}\int_0^1\frac{\Gamma(\delta+r+\nu)}{\Gamma(\delta+r)\Gamma(\nu)}y^{\delta+r-1}(1-y)^{\nu-1}dy =$$
$$\frac{\Gamma(\delta+\nu)\Gamma(\delta+r)}{\Gamma(\delta)\Gamma(\delta+r+\nu)}$$
for $r > -\delta$ since $1 = \int_0^1$ beta$(\delta+r,\nu)$ pdf.

1.25. $E(e^{tY}) = \sum_{y=1}^\infty e^{ty}\frac{-1}{\log(1-\theta)}\frac{1}{y}\exp[\log(\theta)y]$. But $e^{ty}\exp[\log(\theta)y] = \exp[(\log(\theta)+t)y] = \exp[(\log(\theta)+\log(e^t))y] = \exp[\log(\theta e^t)y]$.
So $E(e^{tY}) = \frac{-1}{\log(1-\theta)}[-\log(1-\theta e^t)]\sum_{y=1}^\infty\frac{-1}{\log(1-\theta e^t)}\frac{1}{y}\exp[\log(\theta e^t)y] =$
$\frac{\log(1-\theta e^t)}{\log(1-\theta)}$ since $1 = \sum$ [logarithmic (θe^t) pmf] if $0 < \theta e^t < 1$ or $0 < e^t < 1/\theta$ or $-\infty < t < -\log(\theta)$.

1.28. a) $EX = 0.9EZ + 0.1EW = 0.9\nu\lambda + 0.1(10) = 0.9(3)(4) + 1 = 11.8$.
b) $EX^2 = 0.9[V(Z) + (E(Z))^2] + 0.1[V(W) + (E(W))^2]$
$= 0.9[\nu\lambda^2 + (\nu\lambda)^2] + 0.1[10 + (10)^2]$
$= 0.9[3(16) + 9(16)] + 0.1(110) = 0.9(192) + 11 = 183.8$.

2.8. a) $F_W(w) = P(W \le w) = P(Y \le w-\mu) = F_Y(w-\mu)$. So $f_W(w) = \frac{d}{dw}F_Y(w-\mu) = f_Y(w-\mu)$.

b) $F_W(w) = P(W \leq w) = P(Y \leq w/\sigma) = F_Y(w/\sigma)$. So $f_W(w) = \frac{d}{dw}F_Y(w/\sigma) = f_Y(w/\sigma)\frac{1}{\sigma}$.

c) $F_W(w) = P(W \leq w) = P(\sigma Y \leq w - \mu) = F_Y(\frac{w-\mu}{\sigma})$. So $f_W(w) = \frac{d}{dw}F_Y(\frac{w-\mu}{\sigma}) = f_Y(\frac{w-\mu}{\sigma})\frac{1}{\sigma}$.

2.9. a) See Example 2.16.

2.11. $W = Z^2 \sim \chi^2_1$ where $Z \sim N(0,1)$. So the pdf of W is

$$f(w) = \frac{w^{\frac{1}{2}-1}e^{-\frac{w}{2}}}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} = \frac{1}{\sqrt{w}\sqrt{2\pi}}e^{-\frac{w}{2}}$$

for $w > 0$.

2.12. $(Y-\mu)/\sigma = |Z| \sim HN(0,1)$ where $Z \sim N(0,1)$. So $(Y-\mu)^2 = \sigma^2 Z^2 \sim \sigma^2\chi^2_1 \sim G(0.5, 2\sigma^2)$.

2.16. a) $y = e^{-w} = t^{-1}(w)$, and

$$\left|\frac{dt^{-1}(w)}{dw}\right| = |-e^{-w}| = e^{-w}.$$

Now $P(Y=0) = 0$ so $0 < Y \leq 1$ implies that $W = -\log(Y) > 0$. Hence

$$f_W(w) = f_Y(t^{-1}(w))\left|\frac{dt^{-1}(w)}{dw}\right| = \frac{1}{\lambda}(e^{-w})^{\frac{1}{\lambda}-1}e^{-w} = \frac{1}{\lambda}e^{-w/\lambda}$$

for $w > 0$ which is the EXP(λ) pdf.

2.18. a)

$$f(y) = \frac{1}{\lambda}\frac{\phi y^{\phi-1}}{(1+y^{\phi})^{\frac{1}{\lambda}+1}}$$

where y, ϕ, and λ are all positive. Since $Y > 0$, $W = \log(1+Y^{\phi}) > \log(1) > 0$ and the support $\mathcal{W} = (0,\infty)$. Now $1 + y^{\phi} = e^w$, so $y = (e^w - 1)^{1/\phi} = t^{-1}(w)$. Hence

$$\left|\frac{dt^{-1}(w)}{dw}\right| = \frac{1}{\phi}(e^w - 1)^{\frac{1}{\phi}-1}e^w$$

since $w > 0$. Thus

$$f_W(w) = f_Y(t^{-1}(w))\left|\frac{dt^{-1}(w)}{dw}\right| = \frac{1}{\lambda}\frac{\phi(e^w-1)^{\frac{\phi-1}{\phi}}}{(1+(e^w-1)^{\frac{\phi}{\phi}})^{\frac{1}{\lambda}+1}}\frac{1}{\phi}(e^w-1)^{\frac{1}{\phi}-1}e^w$$

$$= \frac{1}{\lambda} \frac{(e^w - 1)^{1-\frac{1}{\phi}} (e^w - 1)^{\frac{1}{\phi}-1}}{(e^w)^{\frac{1}{\lambda}+1}} e^w$$

$$\frac{1}{\lambda} e^{-w/\lambda}$$

for $w > 0$ which is the EXP(λ) pdf.

2.25. b)

$$f(y) = \frac{1}{\pi\sigma[1 + (\frac{y-\mu}{\sigma})^2]}$$

where y and μ are real numbers and $\sigma > 0$. Now $w = \log(y) = t^{-1}(w)$ and $W = e^Y > 0$ so the support $\mathcal{W} = (0, \infty)$. Thus

$$\left| \frac{dt^{-1}(w)}{dw} \right| = \frac{1}{y},$$

and

$$f_W(w) = f_Y(t^{-1}(w)) \left| \frac{dt^{-1}(w)}{dw} \right| = \frac{1}{\pi\sigma} \frac{1}{[1 + (\frac{\log(y)-\mu}{\sigma})^2]} \frac{1}{y} =$$

$$\frac{1}{\pi\sigma y[1 + (\frac{\log(y)-\mu}{\sigma})^2]}$$

for $y > 0$ which is the $LC(\mu, \sigma)$ pdf.

2.63. a) $EX = E[E[X|Y]] = E[\beta_o + \beta_1 Y] = \beta_0 + 3\beta_1$.

b) $V(X) = E[V(X|Y)] + V[E(X|Y)] = E(Y^2) + V(\beta_0 + \beta_1 Y) = V(Y) + [E(Y)]^2 + \beta_1^2 V(Y) = 10 + 9 + \beta_1^2 10 = 19 + 10\beta_1^2$.

2.64. a) $X_2 \sim N(100, 6)$.

b)

$$\begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 49 \\ 17 \end{pmatrix}, \begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix} \right).$$

c) $X_1 \perp\!\!\!\perp X_4$ and $X_3 \perp\!\!\!\perp X_4$.

d)

$$\rho(X_1, X_2) = \frac{Cov(X_1, X_3)}{\sqrt{\text{VAR}(X_1)\text{VAR}(X_3)}} = \frac{-1}{\sqrt{3}\sqrt{4}} = -0.2887.$$

2.65. a) $Y|X \sim N(49, 16)$ since $Y \perp\!\!\!\perp X$. (Or use $E(Y|X) = \mu_Y + \Sigma_{12}\Sigma_{22}^{-1}(X - \mu_x) = 49 + 0(1/25)(X - 100) = 49$ and $\text{VAR}(Y|X) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = 16 - 0(1/25)0 = 16$.)

b) $E(Y|X) = \mu_Y + \Sigma_{12}\Sigma_{22}^{-1}(X - \mu_x) = 49 + 10(1/25)(X - 100) = 9 + 0.4X$.

c) $\text{VAR}(Y|X) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = 16 - 10(1/25)10 = 16 - 4 = 12$.

2.68. a) $E(Y) = E[E(Y|\Lambda)] = E(\Lambda) = 1$.

b) $V(Y) = E[V(Y|\Lambda)] + V[E(Y|\Lambda)] = E(\Lambda) + V(\Lambda) = 1 + (1)^2 = 2$.

2.71.

y	0	1
$f_{Y_1}(y) = P(Y_1 = y)$	0.76	0.24

So $m(t) = \sum_y e^{ty} f(y) = \sum_y e^{ty} P(Y = y) = e^{t0} 0.76 + e^{t1} 0.24$
$= 0.76 + 0.24e^t$.

2.72. No, $f(x, y) \neq f_X(x) f_Y(y) = \frac{1}{2\pi} \exp[\frac{-1}{2}(x^2 + y^2)]$.

2.73. a) $E(Y) = E[E(Y|P)] = E(kP) = kE(P) = k\frac{\delta}{\delta+\nu} = k4/10 = 0.4k$.

b) $V(Y) = E[V(Y|P)] + V(E(Y|P)] = E[kP(1 - P)] + V(kP) =$
$kE(P) - kE(P^2) + k^2 V(P) =$

$$k\frac{\delta}{\delta + \nu} - n\left[\frac{\delta\nu}{(\delta + \nu)^2(\delta + \nu + 1)} + \left(\frac{\delta}{\delta + \nu}\right)^2\right] + k^2 \frac{\delta\nu}{(\delta + \nu)^2(\delta + \nu + 1)}$$

$= k0.4 - k[0.021818 + 0.16] + k^2 0.021818 = 0.021818k^2 + 0.21818k$.

2.74. a)

y_2	0	1	2
$f_{Y_2}(y_2)$	0.55	0.16	0.29

b) $f(y_1|2) = f(y_1, 2)/f_{Y_2}(2)$ and $f(0, 2)/f_{Y_2}(2) = .24/.29$ while $f(1, 2)/f_{Y_2}(2) = .05/.29$

y_1	0	1
$f_{Y_1\|Y_2}(y_1\|y_2 = 2)$	$24/29 \approx 0.8276$	$5/29 \approx 0.1724$

3.1. a) See Section 10.3.
b) See Section 10.10.
c) See Section 10.35.
d) See Example 3.5.

3.2. a) See Section 10.1.
b) See Section 10.6.
c) See Section 10.13.
d) See Section 10.29.
e) See Section 10.32.

3.3. b) See Section 10.16.
c) See Section 10.25.
d) See Section 10.31.
f) See Section 10.36.
g) See Section 10.41.
h) See Section 10.44.

3.4. a) See Section 10.32.
b) See Section 10.32.
c) See Section 10.13.

3.5. a) See Section 10.4.
b) See Section 10.9.
c) See Section 10.11.
d) See Section 10.24.
h) See Section 10.34.
i) See Section 10.37.
j) See Section 10.43.

4.26.

$$f(x) = \frac{\Gamma(2\theta)}{\Gamma(\theta)\Gamma(\theta)} x^{\theta-1}(1-x)^{\theta-1} = \frac{\Gamma(2\theta)}{\Gamma(\theta)\Gamma(\theta)} \exp[(\theta-1)(\log(x)+\log(1-x))],$$

for $0 < x < 1$, a 1 parameter exponential family. Hence $\sum_{i=1}^n (\log(X_i) + \log(1-X_i))$ is a complete minimal sufficient statistic.

4.27. a) and b)

$$f(x) = \frac{1}{\zeta(\nu)} \exp[-\nu \log(x)] I_{\{1,2,...\}}(x)$$

is a 1 parameter regular exponential family. Hence $\sum_{i=1}^n \log(X_i)$ is a complete minimal sufficient statistic.

c) By the Factorization Theorem, $\boldsymbol{W} = (X_1, ..., X_n)$ is sufficient, but $\boldsymbol{W}$ is not minimal since $\boldsymbol{W}$ is not a function of $\sum_{i=1}^n \log(X_i)$.

5.2. The likelihood function $L(\theta) =$

$$\frac{1}{(2\pi)^n}\exp(\frac{-1}{2}[\sum(x_i-\rho\cos\theta)^2+\sum(y_i-\rho\sin\theta)^2]) =$$

$$\frac{1}{(2\pi)^n}\exp(\frac{-1}{2}[\sum x_i^2-2\rho\cos\theta\sum x_i+\rho^2\cos^2\theta+\sum y_i^2-2\rho\sin\theta\sum y_i+\rho^2\sin^2\theta])$$

$$=\frac{1}{(2\pi)^n}\exp(\frac{-1}{2}[\sum x_i^2+\sum y_i^2+\rho^2])\exp(\rho\cos\theta\sum x_i+\rho\sin\theta\sum y_i).$$

Hence the log likelihood log $L(\theta)$

$$= c+\rho\cos\theta\sum x_i+\rho\sin\theta\sum y_i.$$

The derivative with respect to θ is

$$-\rho\sin\theta\sum x_i+\rho\cos\theta\sum y_i.$$

Setting this derivative to zero gives

$$\rho\sum y_i\cos\theta=\rho\sum x_i\sin\theta$$

or

$$\frac{\sum y_i}{\sum x_i}=\tan\theta.$$

Thus

$$\hat{\theta}=\tan^{-1}(\frac{\sum y_i}{\sum x_i}).$$

Now the boundary points are $\theta = 0$ and $\theta = 2\pi$. Hence $\hat{\theta}_{MLE}$ equals 0, 2π, or $\hat{\theta}$ depending on which value maximizes the likelihood.

5.6. See Section 10.4.

5.7. See Section 10.6.

5.8. See Section 10.9.

5.9. See Section 10.10.

5.10. See Section 10.13.

5.11. See Section 10.16.

5.12. See Section 10.22.

5.13. See Section 10.22.

5.14. See Section 10.24.

5.15. See Section 10.31.

5.16. See Section 10.37.

5.17. See Section 10.43.

5.18. See Section 10.3.

5.19. See Section 10.11.

5.20. See Section 10.41,

5.23. a) The log likelihood is $\log L(\tau) = -\frac{n}{2}\log(2\pi\tau) - \frac{1}{2\tau}\sum_{i=1}^n (X_i - \mu)^2$. The derivative of the log likelihood is equal to $-\frac{n}{2\tau} + \frac{1}{2\tau^2}\sum_{i=1}^n (X_i - \mu)^2$. Setting the derivative equal to 0 and solving for τ gives the MLE $\hat{\tau} = \frac{\sum_{i=1}^n (X_i - \mu)^2}{n}$. Now the likelihood is only defined for $\tau > 0$. As τ goes to 0 or ∞, $\log L(\tau)$ tends to $-\infty$. Since there is only one critical point, $\hat{\tau}$ is the MLE.

b) By the invariance principle, the MLE is $\sqrt{\frac{\sum_{i=1}^n (X_i - \mu)^2}{n}}$.

5.28. This problem is nearly the same as finding the MLE of σ^2 when the data are iid $N(\mu, \sigma^2)$ when μ is known. See Problem 5.23. The MLE in a) is $\sum_{i=1}^n (X_i - \mu)^2/n$. For b) use the invariance principle and take the square root of the answer in a).

5.29. See Example 5.5.

5.30.

$$L(\theta) = \frac{1}{\theta\sqrt{2\pi}} e^{-(x-\theta)^2/2\theta^2}$$

$$ln(L(\theta)) = -ln(\theta) - ln(\sqrt{2\pi}) - (x-\theta)^2/2\theta^2$$

$$\begin{aligned}\frac{dln(L(\theta))}{d\theta} &= \frac{-1}{\theta} + \frac{x-\theta}{\theta^2} + \frac{(x-\theta)^2}{\theta^3} \\ &= \frac{x^2}{\theta^3} - \frac{x}{\theta^2} - \frac{1}{\theta}\end{aligned}$$

by solving for θ,

$$\theta = \frac{x}{2} * (-1 + \sqrt{5}),$$

and

$$\theta = \frac{x}{2} * (-1 - \sqrt{5}).$$

But, $\theta > 0$. Thus, $\hat{\theta} = \frac{x}{2} * (-1 + \sqrt{5})$, when $x > 0$, and $\hat{\theta} = \frac{x}{2} * (-1 - \sqrt{5})$, when $x < 0$.

To check with the second derivative

$$\begin{aligned} \frac{d^2 ln(L(\theta))}{d\theta^2} &= -\frac{2\theta + x}{\theta^3} + \frac{3(\theta^2 + \theta x - x^2)}{\theta^4} \\ &= \frac{\theta^2 + 2\theta x - 3x^2}{\theta^4} \end{aligned}$$

but the sign of the θ^4 is always positive, thus the sign of the second derivative depends on the sign of the numerator. Substitute $\hat{\theta}$ in the numerator and simplify, you get $\frac{x^2}{2}(-5 \pm \sqrt{5})$, which is always negative. Hence by the invariance principle, the MLE of θ^2 is $\hat{\theta}^2$.

5.31. a) For any $\lambda > 0$, the likelihood function

$$L(\sigma, \lambda) = \sigma^{n/\lambda} \ I[x_{(1)} \geq \sigma] \ \frac{1}{\lambda^n} \ \exp\left[-(1 + \frac{1}{\lambda}) \sum_{i=1}^{n} \log(x_i)\right]$$

is maximized by making σ as large as possible. Hence $\hat{\sigma} = X_{(1)}$.

b)

$$L(\hat{\sigma}, \lambda) = \hat{\sigma}^{n/\lambda} \ I[x_{(1)} \geq \hat{\sigma}] \ \frac{1}{\lambda^n} \ \exp\left[-(1 + \frac{1}{\lambda}) \sum_{i=1}^{n} \log(x_i)\right].$$

Hence $\log L(\hat{\sigma}, \lambda) =$

$$\frac{n}{\lambda} \log(\hat{\sigma}) - n \log(\lambda) - (1 + \frac{1}{\lambda}) \sum_{i=1}^{n} \log(x_i).$$

Thus

$$\frac{d}{d\lambda} \log L(\hat{\sigma}, \lambda) = \frac{-n}{\lambda^2} \log(\hat{\sigma}) - \frac{n}{\lambda} + \frac{1}{\lambda^2} \sum_{i=1}^{n} \log(x_i) \stackrel{set}{=} 0,$$

or $-n\log(\hat{\sigma}) + \sum_{i=1}^n \log(x_i) = n\lambda$. So

$$\hat{\lambda} = -\log(\hat{\sigma}) + \frac{\sum_{i=1}^n \log(x_i)}{n} = \frac{\sum_{i=1}^n \log(x_i/\hat{\sigma})}{n}.$$

Now

$$\frac{d^2}{d\lambda^2} \log L(\hat{\sigma}, \lambda) = \frac{2n}{\lambda^3}\log(\hat{\sigma}) + \frac{n}{\lambda^2} - \frac{2}{\lambda^3}\sum_{i=1}^n \log(x_i)\Bigg|_{\lambda=\hat{\lambda}}$$

$$= \frac{n}{\hat{\lambda}^2} - \frac{2}{\hat{\lambda}^3}\sum_{i=1}^n \log(x_i/\hat{\sigma}) = \frac{-n}{\hat{\lambda}^2} < 0.$$

Hence $(\hat{\sigma}, \hat{\lambda})$ is the MLE of (σ, λ).

5.32. a) the likelihood

$$L(\lambda) = c\frac{1}{\lambda^n}\exp\left[-(1+\frac{1}{\lambda})\sum \log(x_i)\right],$$

and the log likelihood

$$\log(L(\lambda)) = d - n\log(\lambda) - (1+\frac{1}{\lambda})\sum \log(x_i).$$

Hence

$$\frac{d}{d\lambda}\log(L(\lambda)) = \frac{-n}{\lambda} + \frac{1}{\lambda^2}\sum \log(x_i) \stackrel{set}{=} 0,$$

or $\sum \log(x_i) = n\lambda$ or

$$\hat{\lambda} = \frac{\sum \log(X_i)}{n}.$$

Notice that

$$\frac{d^2}{d\lambda^2}\log(L(\lambda)) = \frac{n}{\lambda^2} - \frac{2\sum \log(x_i)}{\lambda^3}\Bigg|_{\lambda=\hat{\lambda}} =$$

$$\frac{n}{\hat{\lambda}^2} - \frac{2n\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n}{\hat{\lambda}^2} < 0.$$

Hence $\hat{\lambda}$ is the MLE of λ.

b) By invariance, $\hat{\lambda}^8$ is the MLE of λ^8.

5.33. a) The likelihood

$$L(\theta) = c\ e^{-n2\theta} \exp[\log(2\theta) \sum x_i],$$

and the log likelihood

$$\log(L(\theta)) = d - n2\theta + \log(2\theta) \sum x_i.$$

Hence

$$\frac{d}{d\theta} \log(L(\theta)) = -2n + \frac{2}{2\theta} \sum x_i \stackrel{set}{=} 0,$$

or $\sum x_i = 2n\theta$, or

$$\hat{\theta} = \overline{X}/2.$$

Notice that

$$\frac{d^2}{d\theta^2} \log(L(\theta)) = \frac{-\sum x_i}{\theta^2} < 0$$

unless $\sum x_i = 0$.

b) $(\hat{\theta})^4 = (\overline{X}/2)^4$ by invariance.

5.34. $L(0|\boldsymbol{x}) = 1$ for $0 < x_i < 1$, and $L(1|\boldsymbol{x}) = \prod_{i=1}^n \frac{1}{2\sqrt{x_i}}$ for $0 < x_i < 1$. Thus the MLE is 0 if $1 \geq \prod_{i=1}^n \frac{1}{2\sqrt{x_i}}$ and the MLE is 1 if $1 < \prod_{i=1}^n \frac{1}{2\sqrt{x_i}}$.

5.35. a) Notice that $\theta > 0$ and

$$f(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\theta}} \exp\left(\frac{-(y-\theta)^2}{2\theta}\right).$$

Hence the likelihood

$$L(\theta) = c\ \frac{1}{\theta^{n/2}} \exp\left[\frac{-1}{2\theta} \sum (y_i - \theta)^2\right]$$

and the log likelihood

$$\log(L(\theta)) = d\ -\ \frac{n}{2}\log(\theta) - \frac{1}{2\theta}\sum (y_i-\theta)^2 =$$

$$d\ -\ \frac{n}{2}\log(\theta)\ -\ \frac{1}{2}\sum_{i=1}^n \left(\frac{y_i^2}{\theta}\ -\ \frac{2y_i\theta}{\theta}\ +\ \frac{\theta^2}{\theta}\right)$$

$$= d \;-\; \frac{n}{2}\log(\theta) \;-\; \frac{1}{2}\frac{\sum_{i=1}^n y_i^2}{\theta} \;+\; \sum_{i=1}^n y_i \;-\; \frac{1}{2} n\theta.$$

Thus

$$\frac{d}{d\theta}\log(L(\theta)) = \frac{-n}{2}\frac{1}{\theta} + \frac{1}{2}\sum_{i=1}^n y_i^2 \frac{1}{\theta^2} - \frac{n}{2} \stackrel{set}{=} 0,$$

or

$$\frac{-n}{2}\theta^2 - \frac{n}{2}\theta + \frac{1}{2}\sum_{i=1}^n y_i^2 = 0,$$

or

$$n\theta^2 \;+\; n\theta \;-\; \sum_{i=1}^n y_i^2 = 0. \tag{11.2}$$

Now the quadratic formula states that for $a \neq 0$, the quadratic equation $ay^2 + by + c = 0$ has roots

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Applying the quadratic formula to (11.2) gives

$$\theta = \frac{-n \pm \sqrt{n^2 + 4n\sum_{i=1}^n y_i^2}}{2n}.$$

Since $\theta > 0$, a candidate for the MLE is

$$\hat{\theta} = \frac{-n + \sqrt{n^2 + 4n\sum_{i=1}^n Y_i^2}}{2n} = \frac{-1 + \sqrt{1 + 4\frac{1}{n}\sum_{i=1}^n Y_i^2}}{2}.$$

Since $\hat{\theta}$ satisfies (11.2),

$$n\hat{\theta} - \sum_{i=1}^n y_i^2 = -n\hat{\theta}^2. \tag{11.3}$$

Note that

$$\frac{d^2}{d\theta^2}\log(L(\theta)) = \frac{n}{2\theta^2} - \frac{\sum_{i=1}^n y_i^2}{\theta^3} = \frac{1}{2\theta^3}[n\theta - 2\sum_{i=1}^n y_i^2]\Bigg|_{\theta=\hat{\theta}} =$$

$$\frac{1}{2\hat{\theta}^3}[n\hat{\theta} - \sum_{i=1}^n y_i^2 - \sum_{i=1}^n y_i^2] = \frac{1}{2\hat{\theta}^3}[-n\hat{\theta}^2 - \sum_{i=1}^n y_i^2] < 0$$

by (11.3). Since $L(\theta)$ is continuous with a unique root on $\theta > 0$, $\hat{\theta}$ is the MLE.

5.36. a) $L(\theta) = (\theta - x)^2/3$ for $x - 2 \le \theta \le x + 1$. Since $x = 7$, $L(5) = 4.3$, $L(7) = 0$, and $L(8) = 1/3$. So L is maximized at an endpoint and the MLE $\hat{\theta} = 5$.

b) By invariance the MLE is $h(\hat{\theta}) = h(5) = 10 - e^{-25} \approx 10$.

5.37. a) $L(\lambda) = c \frac{1}{\lambda^n} \exp\left(\frac{-1}{2\lambda^2} \sum_{i=1}^n (e^{x_i} - 1)^2\right)$.
Thus

$$\log(L(\lambda)) = d - n \log(\lambda) - \frac{1}{2\lambda^2} \sum_{i=1}^n (e^{x_i} - 1)^2.$$

Hence

$$\frac{d \log(L(\lambda))}{d\lambda} = \frac{-n}{\lambda} + \frac{1}{\lambda^3} \sum (e^{x_i} - 1)^2 \stackrel{set}{=} 0,$$

or $n\lambda^2 = \sum (e^{x_i} - 1)^2$, or

$$\hat{\lambda} = \frac{\sum (e^{X_i} - 1)^2}{n}.$$

Now

$$\frac{d^2 \log(L(\lambda))}{d\lambda^2} = \frac{n}{\lambda^2} - \frac{3}{\lambda^4} \sum (e^{x_i} - 1)^2 \bigg|_{\lambda = \hat{\lambda}}$$

$$= \frac{n}{\hat{\lambda}^2} - \frac{3n}{\hat{\lambda}^4} \hat{\lambda}^2 = \frac{n}{\hat{\lambda}^2}[1 - 3] < 0.$$

So $\hat{\lambda}$ is the MLE.

5.38. a) The likelihood

$$L(\lambda) = \prod f(x_i) = c \left(\prod \frac{1}{x_i}\right) \frac{1}{\lambda^n} \exp\left[\frac{\sum -(\log x_i)^2}{2\lambda^2}\right],$$

and the log likelihood

$$\log(L(\lambda)) = d - \sum \log(x_i) - n \log(\lambda) - \frac{\sum (\log x_i)^2}{2\lambda^2}.$$

Hence

$$\frac{d}{d\lambda} \log(L(\lambda)) = \frac{-n}{\lambda} + \frac{\sum (\log x_i)^2}{\lambda^3} \stackrel{set}{=} 0,$$

or $\sum(\log x_i)^2 = n\lambda^2$, or
$$\hat{\lambda} = \sqrt{\frac{\sum(\log x_i)^2}{n}}.$$
This solution is unique.

Notice that
$$\frac{d^2}{d\lambda^2}\log(L(\lambda)) = \left.\frac{n}{\lambda^2} - \frac{3\sum(\log x_i)^2}{\lambda^4}\right|_{\lambda=\hat{\lambda}}$$
$$= \frac{n}{\hat{\lambda}^2} - \frac{3n\hat{\lambda}^2}{\hat{\lambda}^4} = \frac{-2n}{\hat{\lambda}^2} < 0.$$

Hence
$$\hat{\lambda} = \sqrt{\frac{\sum(\log X_i)^2}{n}}$$
is the MLE of λ.

b)
$$\hat{\lambda}^2 = \frac{\sum(\log X_i)^2}{n}$$
is the MLE of λ^2 by invariance.

6.7. a) The joint density
$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}}\exp[-\frac{1}{2}\sum(x_i-\mu)^2]$$
$$= \frac{1}{(2\pi)^{n/2}}\exp[-\frac{1}{2}(\sum x_i^2 - 2\mu\sum x_i + n\mu^2)]$$
$$= \frac{1}{(2\pi)^{n/2}}\exp[-\frac{1}{2}\sum x_i^2]\exp[n\mu\overline{x} - \frac{n\mu^2}{2}].$$
Hence by the factorization theorem $\overline{X}$ is a sufficient statistic for μ.

b) $\overline{X}$ is sufficient by a) and complete since the $N(\mu, 1)$ family is a regular one parameter exponential family.

c) $E(I_{-(\infty,t]}(X_1)|\overline{X} = \overline{x}) = P(X_1 \leq t|\overline{X} = \overline{x}) = \Phi(\frac{t-\overline{x}}{\sqrt{1-1/n}})$.

d) By Rao-Blackwell-Lehmann-Scheffe,
$$\Phi(\frac{t-\overline{X}}{\sqrt{1-1/n}})$$

is the UMVUE.

6.14. Note that $\sum X_i \sim G(n, \theta)$. Hence MSE(c) $= Var_\theta(T_n(c)) + [E_\theta T_n(c) - \theta]^2 = c^2 Var_\theta(\sum X_i) + [ncE_\theta X - \theta]^2 = c^2 n\theta^2 + [nc\theta - \theta]^2$.

So
$$\frac{d}{dc} MSE(c) = 2cn\theta^2 + 2[nc\theta - \theta]n\theta.$$
Set this equation to 0 to get $2n\theta^2[c + nc - 1] = 0$ or $c(n+1) = 1$. So $c = 1/(n+1)$.

The second derivative is $2n\theta^2 + 2n^2\theta^2 > 0$ so the function is convex and the local min is in fact global.

6.17. a) Since this is an exponential family, $\log(f(x|\lambda)) = -\log(\lambda) - x/\lambda$ and
$$\frac{\partial}{\partial \lambda} \log(f(x|\lambda)) = \frac{-1}{\lambda} + \frac{x}{\lambda^2}.$$
Hence
$$\frac{\partial^2}{\partial \lambda^2} \log(f(x|\lambda)) = \frac{1}{\lambda^2} - \frac{2x}{\lambda^3}$$
and
$$I_1(\lambda) = -E\left[\frac{\partial}{\partial \lambda} \log(f(x|\lambda))\right] = \frac{-1}{\lambda^2} + \frac{2\lambda}{\lambda^3} = \frac{1}{\lambda^2}.$$

b)
$$FCRLB(\tau(\lambda)) = \frac{[\tau'(\lambda)]^2}{nI_1(\lambda)} = \frac{4\lambda^2}{n/\lambda^2} = 4\lambda^4/n.$$

c) ($T = \sum_{i=1}^n X_i \sim$ Gamma(n, λ) is a complete sufficient statistic. Now $E(T^2) = V(T) + [E(T)]^2 = n\lambda^2 + n^2\lambda^2$. Hence the UMVUE of λ^2 is $T^2/(n + n^2)$.) No, W is a nonlinear function of the complete sufficient statistic T.

6.19.
$$W \equiv S^2(k)/\sigma^2 \sim \chi^2_n/k$$
and
$$MSE(S^2(k)) = MSE(W) = VAR(W) + (E(W) - \sigma^2)^2$$
$$= \frac{\sigma^4}{k^2} 2n + (\frac{\sigma^2 n}{k} - \sigma^2)^2$$
$$= \sigma^4[\frac{2n}{k^2} + (\frac{n}{k} - 1)^2] = \sigma^4 \frac{2n + (n-k)^2}{k^2}.$$

Now the derivative $\frac{d}{dk} MSE(S^2(k))/\sigma^4 =$

$$\frac{-2}{k^3}[2n + (n-k)^2] + \frac{-2(n-k)}{k^2}.$$

Set this derivative equal to zero. Then

$$2k^2 - 2nk = 4n + 2(n-k)^2 = 4n + 2n^2 - 4nk + 2k^2.$$

Hence

$$2nk = 4n + 2n^2$$

or $k = n + 2$.

Should also argue that $k = n+2$ is the global minimizer. Certainly need $k > 0$ and the absolute bias will tend to ∞ as $k \to 0$ and the bias tends to σ^2 as $k \to \infty$, so $k = n+2$ is the unique critical point and is the global minimizer.

6.20. a) Let $W = X^2$. Then $f(w) = f_X(\sqrt{w}) \ 1/(2\sqrt{w}) = (1/\theta)\exp(-w/\theta)$ and $W \sim \exp(\theta)$. Hence $E_\theta(X^2) = E_\theta(W) = \theta$.

b) This is an exponential family and

$$\log(f(x|\theta)) = \log(2x) - \log(\theta) - \frac{1}{\theta}x^2$$

for $x < 0$. Hence

$$\frac{\partial}{\partial\theta} f(x|\theta) = \frac{-1}{\theta} + \frac{1}{\theta^2}x^2$$

and

$$\frac{\partial^2}{\partial\theta^2} f(x|\theta) = \frac{1}{\theta^2} + \frac{-2}{\theta^3}x^2.$$

Hence

$$I_1(\theta) = -E_\theta[\frac{1}{\theta^2} + \frac{-2}{\theta^3}x^2] = \frac{1}{\theta^2}$$

by a). Now

$$CRLB = \frac{[\tau'(\theta)]^2}{nI_1(\theta)} = \frac{\theta^2}{n}$$

where $\tau(\theta) = \theta$.

c) This is a regular exponential family so $\sum_{i=1}^n X_i^2$ is a complete sufficient statistic. Since

$$E_\theta[\frac{\sum_{i=1}^n X_i^2}{n}] = \theta,$$

the UMVUE is $\frac{\sum_{i=1}^n X_i^2}{n}$.

6.21. a) In normal samples, $\overline{X}$ and S are independent, hence

$$Var_\theta[W(\alpha)] = \alpha^2 Var_\theta(T_1) + (1-\alpha)^2 Var_\theta(T_2).$$

b) $W(\alpha)$ is an unbiased estimator of θ. Hence MSE($W(\alpha) \equiv MSE(\alpha) = Var_\theta[W(\alpha)]$ which is found in part a).

c) Now

$$\frac{d}{d\alpha} MSE(\alpha) = 2\alpha Var_\theta(T_1) - 2(1-\alpha) Var_\theta(T_2) = 0.$$

Hence

$$\hat{\alpha} = \frac{Var_\theta(T_2)}{Var_\theta(T_1) + Var_\theta(T_2)} \approx \frac{\frac{\theta^2}{2n}}{\frac{\theta^2}{2n} + \frac{2\theta^2}{2n}} = 1/3$$

using the approximation and the fact that Var($\bar{X}$) $= \theta^2/n$. Note that the second derivative

$$\frac{d^2}{d\alpha^2} MSE(\alpha) = 2[Var_\theta(T_1) + Var_\theta(T_2)] > 0,$$

so $\alpha = 1/3$ is a local min. The critical value was unique, hence 1/3 is the global min.

6.22. a) $X_1 - X_2 \sim N(0, 2\sigma^2)$. Thus,

$$\begin{aligned} E(T_1) &= \int_0^\infty u \frac{1}{\sqrt{4\pi\sigma^2}} e^{\frac{-u^2}{4\sigma^2}} du \\ &= \frac{\sigma}{\sqrt{\pi}}. \end{aligned}$$

$$\begin{aligned} E(T_1^2) &= \frac{1}{2} \int_0^\infty u^2 \frac{1}{\sqrt{4\pi\sigma^2}} e^{\frac{-u^2}{4\sigma^2}} du \\ &= \frac{\sigma^2}{2}. \end{aligned}$$

$V(T_1) = \sigma^2(\frac{1}{2} - \frac{1}{\pi})$ and

$$MSE(T_1) = \sigma^2[(\frac{1}{\sqrt{\pi}}) - 1)^2 + \frac{1}{2} - \frac{1}{\pi}] = \sigma^2[\frac{3}{2} - \frac{2}{\sqrt{\pi}}].$$

b) $\frac{X_i}{\sigma}$ has a N(0,1) and $\frac{\sum_{i=1}^n X_i^2}{\sigma^2}$ has a chi square distribution with n degrees of freedom. Thus

$$E(\sqrt{\frac{\sum_{i=1}^n X_i^2}{\sigma^2}}) = \frac{\sqrt{2}\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})},$$

and

$$E(T_2) = \frac{\sigma}{\sqrt{n}} \frac{\sqrt{2}\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}.$$

Therefore,

$$E(\frac{\sqrt{n}}{\sqrt{2}} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n+1}{2})} T_2) = \sigma.$$

6.23. This is a regular one parameter exponential family with complete sufficient statistic $T_n = \sum_{i=1}^n X_i \sim G(n, \lambda)$. Hence $E(T_n) = n\lambda$, $E(T_n^2) = V(T_n) + (E(T_n))^2 = n\lambda^2 + n^2\lambda^2$, and $T_n^2/(n+n^2)$ is the UMVUE of λ^2.

6.24.

$$\frac{1}{X_i} = \frac{W_i}{\sigma} \sim \frac{\chi_1^2}{\sigma}.$$

Hence if

$$T = \sum_{i=1}^n \frac{1}{X_i}, \text{ then } \mathrm{E}(\frac{\mathrm{T}}{\mathrm{n}}) = \frac{\mathrm{n}}{\mathrm{n}\sigma},$$

and T/n is the UMVUE since $f(x)$ is an exponential family with complete sufficient statistic $1/X$.

6.25. The pdf of T is

$$g(t) = \frac{2nt^{2n-1}}{\theta^{2n}}$$

for $0 < t < \theta$.
$E(T) = \frac{2n}{2n+1}\theta$ and $E(T^2) = \frac{2n}{2n+2}\theta^2$.

$$MSE(CT) = (C\frac{2n}{2n+1}\theta - \theta)^2 + C^2[\frac{2n}{2n+2}\theta^2 - (\frac{2n}{2n+1}\theta)^2]$$

$$\frac{dMSE(CT)}{dC} = 2[\frac{2cn\theta}{2n+1} - \theta][\frac{2n\theta}{2n+1}] + 2c[\frac{2n\theta^2}{2n+2} - \frac{4n^2\theta^2}{(2n+1)^2}].$$

Solve $\frac{dMSE(CT)}{dC} = 0$ to get

$$C = 2\frac{n+1}{2n+1}.$$

Check with the second derivative $\frac{d^2 MSE(CT)}{dC^2} = 4\frac{n\theta^2}{2n+2}$, which is always positive.

6.26. a) $E(Y_i) = 2\theta/3$ and $V(Y_i) = \theta^2/18$. So bias of $T = B(T) = Ec\overline{X} - \theta = c\frac{2}{3}\theta - \theta$ and Var$(T) =$

$$Var(\frac{c\sum X_i}{n}) = \frac{c^2}{n^2}\sum Var(X_i) = \frac{c^2}{n^2}\frac{n\theta^2}{18}.$$

So MSE = Var(T) $+[B(T)]^2 =$

$$\frac{c^2\theta^2}{18n} + (\frac{2\theta}{3}c - \theta)^2.$$

b)

$$\frac{dMSE(c)}{dc} = \frac{2c\theta^2}{18n} + 2(\frac{2\theta}{3}c - \theta)\frac{2\theta}{3}.$$

Set this equation equal to 0 and solve, so

$$\frac{\theta^2 2c}{18n} + \frac{4}{3}\theta(\frac{2}{3}\theta c - \theta) = 0$$

or

$$c[\frac{2\theta^2}{18n} + \frac{8}{9}\theta^2] = \frac{4}{3}\theta^2$$

or

$$c(\frac{1}{9n} + \frac{8}{9}\theta^2) = \frac{4}{3}\theta^2$$

or

$$c(\frac{1}{9n} + \frac{8n}{9n}) = \frac{4}{3}$$

or

$$c = \frac{9n}{1+8n}\frac{4}{3} = \frac{12n}{1+8n}.$$

This is a global min since the MSE is a quadratic in c^2 with a positive coefficient, or because

$$\frac{d^2 MSE(c)}{dc^2} = \frac{2\theta^2}{18n} + \frac{8\theta^2}{9} > 0.$$

6.27. See Example 6.5.

7.6. For both a) and b), the test is reject Ho iff $\prod_{i=1}^n x_i(1-x_i) > c$ where $P_{\theta=1}[\prod_{i=1}^n x_i(1-x_i) > c] = \alpha$.

7.10. H says $f(x) = e^{-x}$ while K says

$$f(x) = x^{\theta-1} e^{-x} / \Gamma(\theta).$$

The monotone likelihood ratio property holds for $\prod x_i$ since then

$$\frac{f_n(\boldsymbol{x}, \theta_2)}{f_n(\boldsymbol{x}, \theta_1)} = \frac{(\prod_{i=1}^n x_i)^{\theta_2 - 1} (\Gamma(\theta_1))^n}{(\prod_{i=1}^n x_i)^{\theta_1 - 1} (\Gamma(\theta_2))^n} = (\frac{\Gamma(\theta_1)}{\Gamma(\theta_2)})^n (\prod_{i=1}^n x_i)^{\theta_2 - \theta_1}$$

which increases as $\prod_{i=1}^n x_i$ increases if $\theta_2 > \theta_1$. Hence the level α UMP test rejects H if

$$\prod_{i=1}^n X_i > c$$

where

$$P_H(\prod_{i=1}^n X_i > c) = P_H(\sum \log(X_i) > \log(c)) = 1 - \alpha.$$

7.11. See Example 7.6.

7.13. Let $\theta_1 = 4$. By Neyman Pearson lemma, reject Ho if

$$\frac{f(\boldsymbol{x}|\theta_1)}{f(\boldsymbol{x}|2)} = \left(\frac{\log(\theta_1)}{\theta - 1}\right)^n \theta_1^{\sum x_i} \left(\frac{1}{\log(2)}\right)^n \frac{1}{2^{\sum x_i}} > k$$

iff

$$\left(\frac{\log(\theta_1)}{(\theta - 1)\log(2)}\right)^n \left(\frac{\theta_1}{2}\right)^{\sum x_i} > k$$

iff

$$\left(\frac{\theta_1}{2}\right)^{\sum x_i} > k'$$

iff

$$\sum x_i \log(\theta_1/2) > c'.$$

So reject Ho iff $\sum X_i > c$ where $P_{\theta=2}(\sum X_i > c) = \alpha$.

7.14. a) By NP lemma reject Ho if

$$\frac{f(\boldsymbol{x}|\sigma = 2)}{f(\boldsymbol{x}|\sigma = 1)} > k'.$$

The LHS =
$$\frac{\frac{1}{2^{3n}}\exp[\frac{-1}{8}\sum x_i^2]}{\exp[\frac{-1}{2}\sum x_i^2]}$$
So reject Ho if
$$\frac{1}{2^{3n}}\exp[\sum x_i^2(\frac{1}{2}-\frac{1}{8})] > k'$$
or if $\sum x_i^2 > k$ where $P_{Ho}(\sum x_i^2 > k) = \alpha$.

b) In the above argument, with any $\sigma_1 > 1$, get
$$\sum x_i^2(\frac{1}{2}-\frac{1}{2\sigma_1^2})$$
and
$$\frac{1}{2}-\frac{1}{2\sigma_1^2} > 0$$
for any $\sigma_1^2 > 1$. Hence the UMP test is the same as in a).

7.15. a) By NP lemma reject Ho if
$$\frac{f(\boldsymbol{x}|\sigma=2)}{f(\boldsymbol{x}|\sigma=1)} > k'.$$
The LHS =
$$\frac{\frac{1}{2^n}\exp[\frac{-1}{8}\sum[\log(x_i)]^2]}{\exp[\frac{-1}{2}\sum[\log(x_i)]^2]}$$
So reject Ho if
$$\frac{1}{2^n}\exp[\sum[\log(x_i)]^2(\frac{1}{2}-\frac{1}{8})] > k'$$
or if $\sum[\log(X_i)]^2 > k$ where $P_{Ho}(\sum[\log(X_i)]^2 > k) = \alpha$.

b) In the above argument, with any $\sigma_1 > 1$, get
$$\sum[\log(x_i)]^2(\frac{1}{2}-\frac{1}{2\sigma_1^2})$$
and
$$\frac{1}{2}-\frac{1}{2\sigma_1^2} > 0$$
for any $\sigma_1^2 > 1$. Hence the UMP test is the same as in a).

7.16. The most powerful test will have the following form

Reject H_0 iff $\frac{f_1(x)}{f_0(x)} > k$.

But $\frac{f_1(x)}{f_0(x)} = 4x^{-\frac{3}{2}}$ and hence we reject H_0 iff X is small, i.e. reject H_0 is $X < k$ for some constant k. This test must also have the size α, that is we require:

$\alpha = P(X < k)$ when $f(x) = f_0(x)) = \int_0^k \frac{3}{64}x^2 dx = \frac{1}{64}k^3$,

so that $k = 4\alpha^{\frac{1}{3}}$.

For the power, when $k = 4\alpha^{\frac{1}{3}}$

$P[X < k$ when $f(x) = f_1(x)] = \int_0^k \frac{3}{16}\sqrt{x}dx = \sqrt{\alpha}$

When $\alpha = 0.01$, the power is $= 0.10$.

8.1 c) The histograms should become more like a normal distribution as n increases from 1 to 200. In particular, when $n = 1$ the histogram should be right skewed while for $n = 200$ the histogram should be nearly symmetric. Also the scale on the horizontal axis should decrease as n increases.

d) Now $\overline{Y} \sim N(0, 1/n)$. Hence the histograms should all be roughly symmetric, but the scale on the horizontal axis should be from about $-3/\sqrt{n}$ to $3/\sqrt{n}$.

8.3. a) $E(X) = \frac{3\theta}{\theta+1}$, thus

$\sqrt{n}(\overline{X} - E(x)) \to N(0, V(x))$, but

$V(x) = \frac{9\theta}{(\theta+2)(\theta+1)^2}$. Let $g(y) = \frac{y}{3-y}$, thus $g'(y) = \frac{3}{(3-y)^2}$. Using delta method

$\sqrt{n}(T_n - \theta) \to N(0, \frac{\theta(\theta+1)^2}{\theta+2})$.

b) It is asymptotically efficient if $\sqrt{n}(T_n - \theta) \to N(0, \nu(\theta))$, where

$$\nu(\theta) = \frac{\frac{d}{d\theta}(\theta)}{-E(\frac{d^2}{d\theta^2} lnf(x|\theta))}$$

But, $E((\frac{d^2}{d\theta^2} lnf(x|\theta)) = \frac{1}{\theta^2}$. Thus $\nu(\theta) = \theta^2 \neq \frac{\theta(\theta+1)^2}{\theta+2}$

c) $\overline{X} \to \frac{3\theta}{\theta+1}$ in probability. Thus $T_n \to \theta$ in probability.

8.5. See Example 8.8.

8.7. a) See Example 8.7.

8.13. a) $Y_n \stackrel{D}{=} \sum_{i=1}^n X_i$ where the X_i are iid χ_1^2. Hence $E(X_i) = 1$ and

$\text{Var}(X_i) = 2$. Thus by the CLT,

$$\sqrt{n}\left(\frac{Y_n}{n} - 1\right) \stackrel{D}{=} \sqrt{n}\left(\frac{\sum_{i=1}^n X_i}{n} - 1\right) \stackrel{D}{\rightarrow} N(0,2).$$

b) Let $g(\theta) = \theta^3$. Then $g'(\theta) = 3\theta^2$, $g'(1) = 3$, and by the delta method,

$$\sqrt{n}\left[\left(\frac{Y_n}{n}\right)^3 - 1\right] \stackrel{D}{\rightarrow} N(0, 2(g'(1))^2) = N(0, 18).$$

8.27. a) See Example 8.1b.

b) See Example 8.3.

8.28. a) By the CLT, $\sqrt{n}(\overline{X} - \lambda)/\sqrt{\lambda} \stackrel{D}{\rightarrow} N(0,1)$. Hence $\sqrt{n}(\overline{X} - \lambda) \stackrel{D}{\rightarrow} N(0, \lambda)$.

b) Let $g(\lambda) = \lambda^3$ so that $g'(\lambda) = 3\lambda^2$ then $\sqrt{n}](\overline{X})^3 - (\lambda)^3] \stackrel{D}{\rightarrow} N(0, \lambda[g'(\lambda)]^2) = N(0, 9\lambda^5)$.

8.29. a) $\overline{X}$ is a complete sufficient statistic. Also, we have $\frac{(n-1)S^2}{\sigma^2}$ has a chi square distribution with $df = n-1$, thus since σ^2 is known the distribution of S^2 does not depend on μ, so S^2 is ancillary. Thus, by Basu's Theorem $\overline{X}$ and S^2 are independent.

b) by CLT (n is large) $\sqrt{n}(\overline{X} - \mu)$ has approximately normal distribution with mean 0 and variance σ^2. Let $g(x) = x^3$, thus, $g^{'}(x) = 3x^2$. Using delta method $\sqrt{n}(g(\overline{X}) - g(\mu))$ goes in distribution to $N(0, \sigma^2(g^{'}(\mu))^2)$ or $\sqrt{n}(\overline{X}^3 - \mu^3)$ goes in distribution to $N(0, \sigma^2(3\mu^2)^2)$. Thus the distribution of $\overline{X}^3$ is approximately normal with mean μ^3 and variance $\frac{9\sigma^2\mu^4}{9}$.

8.30. a) According to the standard theorem, $\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow N(0, 3)$.

b) $E(Y) = \theta, Var(Y) = \frac{\pi^2}{3}$, according to CLT we have $\sqrt{n}(\overline{Y}_n - \theta) \rightarrow N(0, \frac{\pi^2}{3})$.

c) $MED(Y) = \theta$, then $\sqrt{n}(MED(n) - \theta) \rightarrow N(0, \frac{1}{4 f^2(MED(Y))})$ and $f(MED(Y)) = \frac{\exp(-(\theta-\theta))}{[1+\exp(-(\theta-\theta))]^2} = \frac{1}{4}$. Thus $\sqrt{n}(MED(n) - \theta) \rightarrow N(0, \frac{1}{4\frac{1}{16}}) \rightarrow \sqrt{n}(MED(n) - \theta) \rightarrow N(0, 4)$.

d) All three estimators are consistent, but $3 < \frac{\pi^2}{3} < 4$, therefore the estimator $\hat{\theta}_n$ is the best, and the estimator MED(n) is the worst.

9.1. a) $\sum_{i=1}^n X_i^b$ is minimal sufficient for a.

b) It can be shown that $\frac{X^b}{a}$ has an exponential distribution with mean 1. Thus, $\frac{2\sum_{i=1}^n X i^b}{a}$ is distributed χ^2_{2n}. Let $\chi^2_{2n,\alpha/2}$ be the upper $100(\frac{1}{2}\alpha)\%$ point of the chi-square distribution with $2n$ degrees of freedom. Thus, we can write

$$1-\alpha = P(\chi^2_{2n,1-\alpha/2} < \frac{2\sum_{i=1}^n X_i^b}{a} < \chi^2_{2n,\alpha/2})$$

which translates into

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

as a two sided $(1-\alpha)$ confidence interval for a. For $\alpha = 0.05$ and $n = 20$, we have $\chi^2_{2n,\alpha/2} = 34.1696$ and $\chi^2_{2n,1-\alpha/2} = 9.59083$. Thus the confidence interval for a is

$$\left(\frac{\sum_{i=1}^n X_i^b}{17.0848}, \frac{\sum_{i=1}^n X_i^b}{4.795415}\right).$$

9.4. Tables are from simulated data but should be similar to the table below.

```
n    p   ccov   acov
50  .01 .4236 .9914 AC CI better
100 .01 .6704 .9406 AC CI better
150 .01 .8278 .9720 AC CI better
200 .01 .9294 .9098 the CIs are about the same
250 .01 .8160 .8160 the CIs are about the same
300 .01 .9158 .9228 the CIs are about the same
350 .01 .9702 .8312 classical is better
400 .01 .9486 .6692 classical is better
450 .01 .9250 .4080 classical is better
```