CHAPTER 6

Section 6.1

1.

a. We use the sample mean, $\overline{x}$ to estimate the population mean m.

$$\hat{m} = \overline{x} = \frac{\Sigma x_i}{n} = \frac{219.80}{27} = 8.1407$$

b. We use the sample median, $\tilde{x} = 7.7$ (the middle observation when arranged in ascending order).

c. We use the sample standard deviation, $s = \sqrt{s^2} = \sqrt{\dfrac{1860.94 - \frac{(219.8)^2}{27}}{26}} = 1.660$

d. With "success" = observation greater than 10, x = # of successes = 4, and $\hat{p} = \frac{x}{n} = \frac{4}{27} = .1481$

e. We use the sample (std dev)/(mean), or $\dfrac{s}{\overline{x}} = \dfrac{1.660}{8.1407} = .2039$

2.

a. With X = # of T's in the sample, the estimator is $\hat{p} = \frac{X}{n}; x = 10,$ so $\hat{p} = \dfrac{10}{20}, = .50$.

b. Here, X = # in sample without TI graphing calculator, and x = 16, so $\hat{p} = \dfrac{16}{20} = .80$

3.

a. We use the sample mean, $\bar{x} = 1.3481$

b. Because we assume normality, the mean = median, so we also use the sample mean $\bar{x} = 1.3481$. We could also easily use the sample median.

c. We use the 90th percentile of the sample:
$\hat{m} + (1.28)\hat{s} = \bar{x} + 1.28s = 1.3481 + (1.28)(.3385) = 1.7814$.

d. Since we can assume normality,

$$P(X < 1.5) \approx P\left(Z < \frac{1.5 - \bar{x}}{s}\right) = P\left(Z < \frac{1.5 - 1.3481}{.3385}\right) = P(Z < .45) = .6736$$

e. The estimated standard error of $\bar{x} = \frac{\hat{s}}{\sqrt{n}} = \frac{s}{\sqrt{n}} = \frac{.3385}{\sqrt{16}} = .0846$

4.

a. $E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = m_1 - m_2$; $\bar{x} - \bar{y} = 8.141 - 8.575 = .434$

b. $V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(\bar{Y}) = s_{\bar{X}}^2 + s_{\bar{Y}}^2 = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$

$s_{\bar{X}-\bar{Y}} = \sqrt{V(\bar{X} - \bar{Y})} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$; The estimate would be

$$s_{\bar{X}-\bar{Y}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{1.66^2}{27} + \frac{2.104^2}{20}} = .5687.$$

c. $\frac{s_1}{s_2} = \frac{1.660}{2.104} = .7890$

d. $V(X - Y) = V(X) + V(Y) = s_1^2 + s_2^2 = 1.66^2 + 2.104^2 = 7.1824$

5. N = 5,000 T = 1,761,300

$\bar{y} = 374.6$ $\bar{x} = 340.6$ $\bar{d} = 34.0$

$\hat{q}_1 = N\bar{x} = (5{,}000)(340.6) = 1{,}703{,}000$

$\hat{q}_2 = T - N\bar{d} = 1{,}761{,}300 - (5{,}000)(34.0) = 1{,}591{,}300$

$$\hat{q}_3 = T\left(\frac{\bar{x}}{\bar{y}}\right) = 1{,}761{,}300\left(\frac{340.6}{374.6}\right) = 1{,}601{,}438.281$$

6.

a. Let $y_i = \ln(x_i)$ for I = 1, .., 31. It is easily verified that the sample mean and sample sd of the $y_i's$ are $\bar{y} = 5.102$ and $s_y = .4961$. Using the sample mean and sample sd to estimate m and s, respectively, gives $\hat{m} = 5.102$ and $\hat{s} = .4961$ (whence $\hat{s}^2 = s_y^2 = .2461$).

b. $E(X) \equiv \exp\left[m + \frac{s^2}{2}\right]$. It is natural to estimate E(X) by using $\hat{m}$ and $\hat{s}^2$ in place of m and s^2 in this expression:

$$E(\hat{X}) = \exp\left[5.102 + \frac{.2461}{2}\right] = \exp(5.225) = 185.87$$

7.

a. $\hat{m} = \bar{x} = \frac{\sum x_i}{n} = \frac{1206}{10} = 120.6$

b. $\hat{t} = 10{,}000 \qquad \hat{m} = 1{,}206{,}000$

c. 8 of 10 houses in the sample used at least 100 therms (the "successes"), so $\hat{p} = \frac{8}{10} = .80.$

d. The ordered sample values are 89, 99, 103, 109, 118, 122, 125, 138, 147, 156, from which the two middle values are 118 and 122, so $\hat{\tilde{m}} = \tilde{x} = \frac{118 + 122}{2} = 120.0$

8.

a. With q denoting the true proportion of defective components,

$$\hat{q} = \frac{(\#\, defective\, in.sample)}{sample.size} = \frac{12}{80} = .150$$

b. P(system works) = p^2, so an estimate of this probability is $\hat{p}^2 = \left(\frac{68}{80}\right)^2 = .723$

9.

a. $E(\overline{X}) = m = E(X) = l$, so $\overline{X}$ is an unbiased estimator for the Poisson parameter l ; $\sum x_i = (0)(18) + (1)(37) + \ldots + (7)(1) = 317$, since n = 150,
$\hat{l} = \overline{x} = \frac{317}{150} = 2.11$.

b. $s_{\overline{x}} = \frac{s}{\sqrt{n}} = \frac{\sqrt{l}}{\sqrt{n}}$, so the estimated standard error is $\sqrt{\frac{\hat{l}}{n}} = \frac{\sqrt{2.11}}{\sqrt{150}} = .119$

10.

a. $E(\overline{X}^2) = Var(\overline{X}) + [E(\overline{X})]^2 = \frac{s^2}{n} + m^2$, so the bias of the estimator $\overline{X}^2$ is $\frac{s^2}{n}$; thus $\overline{X}^2$ tends to overestimate m^2.

b. $E(\overline{X}^2 - kS^2) = E(\overline{X}^2) - kE(S^2) = m^2 + \frac{s^2}{n} - ks^2$, so with $k = \frac{1}{n}$,
$E(\overline{X}^2 - kS^2) = m^2$.

11.

a. $E\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) = \frac{1}{n_1}E(X_1) - \frac{1}{n_2}E(X_2) = \frac{1}{n_1}(n_1 p_1) - \frac{1}{n_2}(n_2 p_2) = p_1 - p_2$.

b. $Var\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) = Var\left(\frac{X_1}{n_1}\right) + Var\left(\frac{X_2}{n_2}\right) = \left(\frac{1}{n_1}\right)^2 Var(X_1) + \left(\frac{1}{n_2}\right)^2 Var(X_2)$
$\frac{1}{n_1^2}(n_1 p_1 q_1) + \frac{1}{n_2^2}(n_2 p_2 q_2) = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$, and the standard error is the square root of this quantity.

c. With $\hat{p}_1 = \frac{x_1}{n_1}$, $\hat{q}_1 = 1 - \hat{p}_1$, $\hat{p}_2 = \frac{x_2}{n_2}$, $\hat{q}_2 = 1 - \hat{p}_2$, the estimated standard error is
$\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$.

d. $(\hat{p}_1 - \hat{p}_2) = \frac{127}{200} - \frac{176}{200} = .635 - .880 = -.245$

e. $\sqrt{\dfrac{(.635)(.365)}{200}+\dfrac{(.880)(.120)}{200}}=.041$

12. $$E\left[\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\right]=\frac{(n_1-1)}{n_1+n_2-2}E(S_1^2)+\frac{(n_2-1)}{n_1+n_2-2}E(S_2^2)$$
$$=\frac{(n_1-1)}{n_1+n_2-2}s^2+\frac{(n_2-1)}{n_1+n_2-2}s^2=s^2.$$

13. $$E(X)=\int_{-1}^{1}x\cdot\tfrac{1}{2}(1+qx)dx=\left.\frac{x^2}{4}+\frac{qx^3}{6}\right|_{-1}^{1}=\frac{1}{3}q \qquad E(X)=\frac{1}{3}q$$
$$E(\overline{X})=\frac{1}{3}q \qquad \hat{q}=3\overline{X}\Rightarrow E(\hat{q})=E(3\overline{X})=3E(\overline{X})=3\left(\frac{1}{3}\right)q=q$$

14.

a. min(x_i) = 202 and max(x_i) = 525, so the estimate of the number of planes manufactured is max(x_i) - min(x_i) + 1 = 525 – 202 + 1 = 324.

b. The estimate will equal the true number of planes manufactured iff min(x_i) = α and max(x_i) = β, i.e., iff the smallest serial number in the population and the largest serial number in the population both appear in the sample. The estimator is not unbiased. This is because max(x_i) never overestimates β and will usually underestimate it (unless max(x_i) = β) , so that E[max(x_i)] < β. Similarly, E[min(x_i)] > α ,so E[max(x_i) - min(x_i)] < β - α + 1; The estimate will usually be smaller than β - α + 1, and can never exceed it.

15.

a. $E(X^2)=2q$ implies that $E\left(\dfrac{X^2}{2}\right)=q$. Consider $\hat{q}=\dfrac{\sum X_i^2}{2n}$. Then
$$E(\hat{q})=E\left(\frac{\sum X_i^2}{2n}\right)=\frac{\sum E(X_i^2)}{2n}=\frac{\sum 2q}{2n}=\frac{2nq}{2n}=q$$, implying that $\hat{q}$ is an unbiased estimator for q .

b. $\sum x_i^2=1490.1058$, so $\hat{q}=\dfrac{1490.1058}{20}=74.505$

16.

a. $E\left[d\overline{X} + (1-d)\overline{Y}\right] = dE(\overline{X}) + (1-d)E(\overline{Y}) = dm + (1-d)m = m$

b. $Var\left[d\overline{X} + (1-d)\overline{Y}\right] = d^2 Var(\overline{X}) + (1-d)^2 Var(\overline{Y}) = \dfrac{d^2 s^2}{m} + \dfrac{4(1-d)^2 s^2}{n}$.

Setting the derivative with respect to d equal to 0 yields $\dfrac{2d s^2}{m} + \dfrac{8(1-d)s^2}{n} = 0$,

from which $d = \dfrac{4m}{4m+n}$.

17.

a.
$$E(\hat{p}) = \sum_{x=0}^{\infty} \frac{r-1}{x+r-1} \cdot \binom{x+r-1}{x} \cdot p^r \cdot (1-p)^x$$
$$= p\sum_{x=0}^{\infty} \frac{(x+r-2)!}{x!(r-2)!} \cdot p^{r-1} \cdot (1-p)^x = p\sum_{x=0}^{\infty} \binom{x+r-2}{x} p^{r-1}(1-p)^x$$
$$= p\sum_{x=0}^{\infty} nb(x; r-1, p) = p.$$

b. For the given sequence, x = 5, so $\hat{p} = \dfrac{5-1}{5+5-1} = \dfrac{4}{9} = .444$

18.

a. $f(x; m, s^2) = \dfrac{1}{\sqrt{2p}\, s} e^{-\left((x-m)^2 / 2s^2\right)}$, so $f(m; m, s^2) = \dfrac{1}{\sqrt{2p}\, s}$ and

$\dfrac{1}{4n[[f(m)]^2} = \dfrac{2p s^2}{4n} = \dfrac{p}{2} \cdot \dfrac{s^2}{n}$; since $\dfrac{p}{2} > 1$, $Var(\widetilde{X}) > Var(\overline{X})$.

b. $f(m) = \dfrac{1}{p}$, so $Var(\widetilde{X}) \approx \dfrac{p^2}{4n} = \dfrac{2.467}{n}$.

19.

a. $l = .5p + .15 \Rightarrow 2l = p + .3$, so $p = 2l - .3$ and $\hat{p} = 2\hat{l} - .3 = 2\left(\frac{Y}{n}\right) - .3$; the estimate is $2\left(\frac{20}{80}\right) - .3 = .2$.

b. $E(\hat{p}) = E\left(2\hat{l} - .3\right) = 2E\left(\hat{l}\right) - .3 = 2l - .3 = p$, as desired.

c. Here $l = .7p + (.3)(.3)$, so $p = \frac{10}{7}l - \frac{9}{70}$ and $\hat{p} = \frac{10}{7}\left(\frac{Y}{n}\right) - \frac{9}{70}$.

Section 6.2

20.

a. We wish to take the derivative of $\ln\left[\binom{n}{x}p^x(1-p)^{n-x}\right]$, set it equal to zero and solve for p. $\frac{d}{dp}\left[\ln\binom{n}{x} + x\ln(p) + (n-x)\ln(1-p)\right] = \frac{x}{p} - \frac{n-x}{1-p}$; setting this equal to zero and solving for p yields $\hat{p} = \frac{x}{n}$. For n = 20 and x = 3, $\hat{p} = \frac{3}{20} = .15$

b. $E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}(np) = p$; thus $\hat{p}$ is an unbiased estimator of p.

c. $(1-.15)^5 = .4437$

21.

a. $E(X) = b \cdot \Gamma\left(1+\frac{1}{a}\right)$ and $E(X^2) = Var(X) + [E(X)]^2 = b^2\Gamma\left(1+\frac{2}{a}\right)$, so the moment estimators $\hat{a}$ and $\hat{b}$ are the solution to $\bar{X} = \hat{b} \cdot \Gamma\left(1+\frac{1}{\hat{a}}\right)$, $\frac{1}{n}\sum X_i^2 = \hat{b}^2\Gamma\left(1+\frac{2}{\hat{a}}\right)$. Thus $\hat{b} = \frac{\bar{X}}{\Gamma\left(1+\frac{1}{\hat{a}}\right)}$, so once $\hat{a}$ has been determined $\Gamma\left(1+\frac{1}{\hat{a}}\right)$ is evaluated and $\hat{b}$ then computed. Since $\bar{X}^2 = \hat{b}^2 \cdot \Gamma^2\left(1+\frac{1}{\hat{a}}\right)$,

$$\frac{1}{n}\sum \frac{X_i^2}{\bar{X}^2} = \frac{\Gamma\left(1+\frac{2}{\hat{a}}\right)}{\Gamma^2\left(1+\frac{1}{\hat{a}}\right)}$$, so this equation must be solved to obtain $\hat{a}$.

b. From **a**, $\frac{1}{20}\left(\frac{16{,}500}{28.0^2}\right) = 1.05 = \frac{\Gamma\left(1+\frac{2}{\hat{a}}\right)}{\Gamma^2\left(1+\frac{1}{\hat{a}}\right)}$, so $\frac{1}{1.05} = .95 = \frac{\Gamma^2\left(1+\frac{1}{\hat{a}}\right)}{\Gamma\left(1+\frac{2}{\hat{a}}\right)}$, and from the hint, $\frac{1}{\hat{a}} = .2 \Rightarrow \hat{a} = 5$. Then $\hat{b} = \frac{\bar{x}}{\Gamma(1.2)} = \frac{28.0}{\Gamma(1.2)}$.

22.

a. $E(X) = \int_0^1 x(q+1)x^q\,dx = \frac{q+1}{q+2} = 1 - \frac{1}{q+2}$, so the moment estimator $\hat{q}$ is the solution to $\bar{X} = 1 - \frac{1}{\hat{q}+2}$, yielding $\hat{q} = \frac{1}{1-\bar{X}} - 2$. Since $\bar{x} = .80, \hat{q} = 5 - 2 = 3.$

b. $f(x_1,\dots,x_n;q) = (q+1)^n (x_1x_2\dots x_n)^q$, so the log likelihood is $n\ln(q+1) + q\sum \ln(x_i)$. Taking $\frac{d}{dq}$ and equating to 0 yields $\frac{n}{q+1} = -\sum \ln(x_i)$, so $\hat{q} = -\frac{n}{\sum \ln(X_i)} - 1$. Taking $\ln(x_i)$ for each given x_i yields ultimately $\hat{q} = 3.12$.

23. For a single sample from a Poisson distribution,

$$f(x_1,\ldots,x_n; l) = \frac{e^{-l}\, l^{x_1}}{x_1!}\cdots\frac{e^{-l}\, l^{x_n}}{x_n!} = \frac{e^{-nl}\, l^{\sum x_1}}{x_1!\ldots x_n!}$$, so

$\ln[f(x_1,\ldots,x_n; l)] = -nl + \sum x_i \ln(l) - \sum \ln(x_i!)$. Thus

$\frac{d}{dl}[\ln[f(x_1,\ldots,x_n; l)]] = -n + \frac{\sum x_i}{l} = 0 \Rightarrow \hat{l} = \frac{\sum x_i}{n} = \bar{x}$. For our problem,

$f(x_1,\ldots,x_n, y_1 \ldots y_n; l_1, l_2)$ is a product of the x sample likelihood and the y sample likelihood, implying that $\hat{l}_1 = \bar{x}, \hat{l}_2 = \bar{y}$, and (by the invariance principle)

$\widehat{(l_1 - l_2)} = \bar{x} - \bar{y}$.

24. We wish to take the derivative of $\ln\left[\binom{x+r-1}{x} p^r (1-p)^x\right]$ with respect to p, set it equal to zero, and solve for p: $\frac{d}{dp}\left[\ln\binom{x+r-1}{x} + r\ln(p) + x\ln(1-p)\right] = \frac{r}{p} - \frac{x}{1-p}$.

Setting this equal to zero and solving for p yields $\hat{p} = \frac{r}{r+x}$. This is the number of successes over the total number of trials, which is the same estimator for the binomial in exercise 6.20. The unbiased estimator from exercise 6.17 is $\hat{p} = \frac{r-1}{r+x-1}$, which is not the same as the maximum likelihood estimator.

25.

a. $\hat{m} = \bar{x} = 384.4; s^2 = 395.16$, so $\frac{1}{n}\sum(x_i - \bar{x})^2 = \hat{s}^2 = \frac{9}{10}(395.16) = 355.64$ and $\hat{s} = \sqrt{355.64} = 18.86$ (this is not s).

b. The 95th percentile is $m + 1.645s$, so the mle of this is (by the invariance principle) $\hat{m} + 1.645\hat{s} = 415.42$.

26. The mle of $P(X \le 400)$ is (by the invariance principle)

$$\Phi\left(\frac{400 - \hat{m}}{\hat{s}}\right) = \Phi\left(\frac{400 - 384.4}{18.86}\right) = \Phi(.80) = .7881$$

27.

a. $f(x_1,\ldots,x_n;a,b)=\dfrac{(x_1x_2\ldots x_n)^{a-1}e^{-\Sigma x_i/b}}{b^{na}\Gamma^n(a)}$, so the log likelihood is

$(a-1)\sum\ln(x_i)-\dfrac{\sum x_i}{b}-na\ln(b)-n\ln\Gamma(a)$. Equating both $\dfrac{d}{da}$ and $\dfrac{d}{db}$ to 0 yields $\sum\ln(x_i)-n\ln(b)-n\dfrac{d}{da}\Gamma(a)=0$ and $\dfrac{\sum x_i}{b^2}=\dfrac{na}{b}=0$, a very difficult system of equations to solve.

b. From the second equation in **a**, $\dfrac{\sum x_i}{b}=na\Rightarrow\bar{x}=ab=m$, so the mle of m is $\hat{m}=\bar{X}$.

28.

a. $\left(\dfrac{x_1}{q}\exp\left[-x_1^2/2q\right]\right)\ldots\left(\dfrac{x_n}{q}\exp\left[-x_n^2/2q\right]\right)=(x_1\ldots x_n)\dfrac{\exp\left[-\Sigma x_i^2/2q\right]}{q^n}$. The natural log of the likelihood function is $\ln(x_i\ldots x_n)-n\ln(q)-\dfrac{\Sigma x_i^2}{2q}$. Taking the derivative wrt q and equating to 0 gives $-\dfrac{n}{q}+\dfrac{\Sigma x_i^2}{2q^2}=0$, so $nq=\dfrac{\Sigma x_i^2}{2}$ and $q=\dfrac{\Sigma x_i^2}{2n}$. The mle is therefore $\hat{q}=\dfrac{\Sigma X_i^2}{2n}$, which is identical to the unbiased estimator suggested in Exercise 15.

b. For x > 0 the cdf of X if $F(x;q)=P(X\le x)$ is equal to $1-\exp\left[\dfrac{-x^2}{2q}\right]$. Equating this to .5 and solving for x gives the median in terms of q: $.5=\exp\left[\dfrac{-x^2}{2q}\right]$ implies that $\ln(.5)=\dfrac{-x^2}{2q}$, so $x=\tilde{m}=\sqrt{1.38630}$. The mle of $\tilde{m}$ is therefore $(1.38630\hat{q})^{\frac{1}{2}}$.

29.

a. The joint pdf (likelihood function) is

$$f(x_1,\dots,x_n; l, q) = \begin{cases} l^n e^{-l\Sigma(x_i - q)} & x_1 \ge q, \dots, x_n \ge q \\ 0 & otherwise \end{cases}$$

Notice that $x_1 \ge q, \dots, x_n \ge q$ iff $\min(x_i) \ge q$,

and that $-l\Sigma(x_i - q) = -l\Sigma x_i + nlq$.

Thus likelihood = $\begin{cases} l^n \exp(-l\Sigma x_i)\exp(nlq) & \min(x_i) \ge q \\ 0 & \min(x_i) < q \end{cases}$

Consider maximization wrt q. Because the exponent nlq is positive, increasing q will increase the likelihood provided that $\min(x_i) \ge q$; if we make q larger than $\min(x_i)$, the likelihood drops to 0. This implies that the mle of q is $\hat{q} = \min(x_i)$. The log likelihood is now $n\ln(l) - l\Sigma(x_i - \hat{q})$. Equating the derivative wrt l to 0 and solving yields $\hat{l} = \dfrac{n}{\Sigma(x_i - \hat{q})} = \dfrac{n}{\Sigma x_i - n\hat{q}}$.

b. $\hat{q} = \min(x_i) = .64$, and $\Sigma x_i = 55.80$, so $\hat{l} = \dfrac{10}{55.80 - 6.4} = .202$

30. The likelihood is $f(y; n, p) = \binom{n}{y} p^y (1-p)^{n-y}$ where

$p = P(X \ge 24) = 1 - \int_0^{24} l e^{-lx}dx = e^{-24l}$. We know $\hat{p} = \dfrac{y}{n}$, so by the invariance principle $e^{-24l} = \dfrac{y}{n} \Rightarrow \hat{l} = -\dfrac{\left[\ln\left(\frac{y}{n}\right)\right]}{24} = .0120$ for n = 20, y = 15.

Supplementary Exercises

31. $P\left(\left|\overline{X} - m\right| > e\right) = P\left(\overline{X} - m > e\right) + P\left(\overline{X} - m < -e\right) = P\left(\dfrac{\overline{X} - m}{s/\sqrt{n}} > \dfrac{e}{s/\sqrt{n}}\right) + P\left(\dfrac{\overline{X} - m}{s/\sqrt{n}} < \dfrac{-e}{s/\sqrt{n}}\right)$

$$= P\left(Z > \frac{\sqrt{n}e}{s}\right) + P\left(Z < \frac{-\sqrt{n}e}{s}\right) = \int_{\sqrt{n}e/s}^{\infty} \frac{1}{\sqrt{2p}} e^{-z^2/2}dz + \int_{-\infty}^{-\sqrt{n}e/s} \frac{1}{\sqrt{2p}} e^{-z^2/2}dz.$$

As $n \to \infty$, both integrals $\to 0$ since $\lim_{c\to\infty} \int_c^{\infty} \frac{1}{\sqrt{2p}} e^{-z^2/2}dz = 0$.

32. sp

a. $F_Y(y) = P(Y \le y) = P(X_1 \le y, \ldots, X_n \le y) = P(X_1 \le y) \ldots P(X_n \le y) = \left(\dfrac{y}{q}\right)^n$

for $0 \le y \le q$, so $f_Y(y) = \dfrac{ny^{n-1}}{q^n}$.

b. $E(Y) = \int_0^q y \cdot \dfrac{ny^{n-1}}{n} dy = \dfrac{n}{n+1} q$. While $\hat{q} = Y$ is not unbiased, $\dfrac{n+1}{n} Y$ is, since $E\left[\dfrac{n+1}{n} Y\right] = \dfrac{n+1}{n} E(Y) = \dfrac{n+1}{n} \cdot \dfrac{n}{n+1} q = q$, so $K = \dfrac{n+1}{n}$ does the trick.

33. Let x_1 = the time until the first birth, x_2 = the elapsed time between the first and second births, and so on. Then $f(x_1, \ldots, x_n; l) = l e^{-l x_1} \cdot (2l) e^{-2 l x_2} \ldots (nl) e^{-n l x_n} = n!\, l^n e^{-l \Sigma k x_k}$. Thus the log likelihood is $\ln(n!) + n \ln(l) - l \Sigma k x_k$. Taking $\dfrac{d}{dl}$ and equating to 0 yields

$\hat{l} = \dfrac{n}{\sum_{k=1}^{n} k x_k}$. For the given sample, n = 6, x_1 = 25.2, x_2 = 41.7 – 25.2 = 16.5, x_3 = 9.5, x_4 = 4.3, x_5 = 4.0, x_6 = 2.3; so $\sum_{k=1}^{6} k x_k = (1)(25.2) + (2)(16.5) + \ldots + (6)(2.3) = 137.7$ and

$\hat{l} = \dfrac{6}{137.7} = .0436$.

34. $MSE(KS^2) = Var(KS^2) + Bias(KS^2)$.

$Bias(KS^2) = E(KS^2) - s^2 = Ks^2 - s^2 = s^2(K-1)$, and

$Var(KS^2) = K^2 Var(S^2) = K^2\left(E\left[(S^2)^2\right] - \left[E(S^2)\right]^2\right) = K^2\left(\dfrac{(n+1)s^4}{n-1} - \left(s^2\right)^2\right)$

$= \left[\dfrac{2K^2}{n-1} + (k-1)^2\right] s^4$. To find the minimizing value of K, take $\dfrac{d}{dK}$ and equate to 0; the result is $K = \dfrac{n-1}{n+1}$; thus the estimator which minimizes MSE is neither the unbiased estimator (K = 1) nor the mle $K = \dfrac{n-1}{n}$.

35.

$x_i + x_j$	23.5	26.3	28.0	28.2	29.4	29.5	30.6	31.6	33.9	49.3
23.5	23.5	24.9	25.75	25.85	26.45	26.5	27.05	27.55	28.7	36.4
26.3		26.3	27.15	27.25	27.85	27.9	28.45	28.95	30.1	37.8
28.0			28.0	28.1	28.7	28.75	29.3	29.8	30.95	38.65
28.2				28.2	28.8	28.85	29.4	29.9	31.05	38.75
29.4					29.4	29.45	30.0	30.5	30.65	39.35
29.5						29.5	30.05	30.55	31.7	39.4
30.6							30.6	31.1	32.25	39.95
31.6								31.6	32.75	40.45
33.9									33.9	41.6
49.3										49.3

There are 55 averages, so the median is the 28$^{\text{th}}$ in order of increasing magnitude. Therefore, $\hat{m} = 29.5$

36. With $\sum x = 555.86$ and $\sum x^2 = 15{,}490$, $s = \sqrt{s^2} = \sqrt{2.1570} = 1.4687$. The $|x_i - \tilde{x}|$'s are, in increasing order, .02, .02, .08, .22, .32, .42, .53, .54, .65, .81, .91, 1.15, 1.17, 1.30, 1.54, 1.54, 1.71, 2.35, 2.92, 3.50. The median of these values is $\frac{(.81 + .91)}{2} = .86$. The estimate based on the resistant estimator is then $\frac{.86}{.6745} = 1.275$. This estimate is in reasonably close agreement with s.

37. Let $c = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \cdot \sqrt{\frac{2}{n-1}}}$. Then E(cS) = cE(S), and c cancels with the two Γ factors and the square root in E(S), leaving just s. When n = 20, $c = \frac{\Gamma(9.5)}{\Gamma(10) \cdot \sqrt{\frac{2}{19}}}$. $\Gamma(10) = 9!$ and $\Gamma(9.5) = (8.5)(7.5)\ldots(1.5)(.5)\Gamma(.5)$, but $\Gamma(.5) = \sqrt{p}$. Straightforward calculation gives c = 1.0132.

38.

a. The likelihood is

$$\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x_i-\mu_i)}{2\sigma^2}}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y_i-\mu_i)}{2\sigma^2}}=\frac{1}{(2\pi\sigma^2)^n}e^{-\frac{\left(\Sigma(x_i-\mu_i)^2+\Sigma(y_i-\mu_i)^2\right)}{2\sigma^2}}$$

. The log likelihood is thus $-n\ln\left(2\pi\sigma^2\right)-\frac{\left(\Sigma(x_i-\mu_i)^2+\Sigma(y_i-\mu_i)^2\right)}{2\sigma^2}$. Taking $\frac{d}{d\mu_i}$ and equating to zero gives $\hat{\mu}_i=\frac{x_i+y_i}{2}$. Substituting these estimates of the $\hat{\mu}_i$'s into the log likelihood gives

$$-n\ln\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left(\sum\left(x_i-\frac{x_i+y_i}{2}\right)^2+\sum\left(y_i-\frac{x_i+y_i}{2}\right)^2\right)$$

$=-n\ln\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left(\frac{1}{2}\Sigma(x_i-y_i)^2\right)$. Now taking $\frac{d}{d\sigma^2}$, equating to zero, and solving for σ^2 gives the desired result.

b. $E(\hat{\sigma})=\frac{1}{4n}E\left(\Sigma(X_i-Y_i)^2\right)=\frac{1}{4n}\cdot\Sigma E(X_i-Y)^2$, but $E(X_i-Y)^2=V(X_i-Y)+\left[E(X_i-Y)\right]^2=2\sigma^2+0=2\sigma^2$. Thus $E(\hat{\sigma}^2)=\frac{1}{4n}\Sigma\left(2\sigma^2\right)=\frac{1}{4n}2n\sigma^2=\frac{\sigma^2}{2}$, so the mle is definitely not unbiased; the expected value of the estimator is only half the value of what is being estimated!