CHAPTER 7

Section 7-2

7-1. The proportion of arrivals for chest pain is 8 among 103 total arrivals. The proportion = 8/103.

7-2. The proportion is 10/80 =1/8.

7-3. $$P(1.009 \le \overline{X} \le 1.012) = P\left(\tfrac{1.009-1.01}{0.003/\sqrt{9}} \le \tfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \tfrac{1.012-1.01}{0.003/\sqrt{9}}\right)$$
$$= P(-1 \le Z \le 2) = P(Z \le 2) - P(Z \le -1) = 0.9772 - 0.1586 = 0.8186$$

7-4. $X_i \sim N(100, 10^2)$ $\quad n = 25$

$$\mu_{\overline{X}} = 100 \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$$
$$P[(100 - 1.8(2)) \le \overline{X} \le (100 + 2)] = P(96.4 \le \overline{X} \le 102) = P\left(\tfrac{96.4-100}{2} \le \tfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \tfrac{102-100}{2}\right)$$
$$= P(-1.8 \le Z \le 1) = P(Z \le 1) - P(Z \le -1.8) = 0.8413 - 0.0359 = 0.8054$$

7-5. $$\mu_{\overline{X}} = 75.5\,psi \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{6}} = 1.429$$
$$P(\overline{X} \ge 75.75) = P\left(\tfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \ge \tfrac{75.75-75.5}{1.429}\right)$$
$$= P(Z \ge 0.175) = 1 - P(Z \le 0.175)$$
$$= 1 - 0.56945 = 0.43055$$

7-6.

n = 6	n = 49
$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{6}} = 1.429$	$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{49}} = 0.5$
$\sigma_{\overline{X}}$ is reduced by 0.929 psi	

7-7. Assuming a normal distribution,

$$\mu_{\overline{X}} = 2500 \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{50}{\sqrt{5}} = 22.361$$
$$P(2499 \le \overline{X} \le 2510) = P\left(\tfrac{2499-2500}{22.361} \le \tfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \tfrac{2510-2500}{22.361}\right)$$
$$= P(-0.045 \le Z \le 0.45) = P(Z \le 0.45) - P(Z \le -0.045)$$
$$= 0.6736 - 0.482 = 0.191$$

7-8. $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{50}{\sqrt{5}} = 22.361\,psi$ = standard error of $\overline{X}$

7-9. $\sigma^2 = 25$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

$$n = \left(\frac{\sigma}{\sigma_{\overline{X}}}\right)^2 = \left(\frac{5}{1.5}\right)^2 = 11.11 \sim 12$$

7-10. Let $Y = \overline{X} - 6$

$$\mu_X = \frac{a+b}{2} = \frac{(0+1)}{2} = \tfrac{1}{2}$$

$$\mu_{\overline{X}} = \mu_X$$

$$\sigma_X^2 = \frac{(b-a)^2}{12} = \tfrac{1}{12}$$

$$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} = \frac{\tfrac{1}{12}}{12} = \frac{1}{144}$$

$$\sigma_{\overline{X}} = \tfrac{1}{12}$$

$$\mu_Y = \tfrac{1}{2} - 6 = -5\tfrac{1}{2}$$

$$\sigma_Y^2 = \tfrac{1}{144}$$

$Y = \overline{X} - 6 \sim N(-5\tfrac{1}{2}, \tfrac{1}{144})$, approximately, using the central limit theorem.

7-11. $n = 36$

$$\mu_X = \frac{a+b}{2} = \frac{(3+1)}{2} = 2$$

$$\sigma_X = \sqrt{\frac{1^2 + 0^2 + 1^2}{3}} = \sqrt{\frac{2}{3}}$$

$$\mu_{\overline{X}} = 2, \sigma_{\overline{X}} = \frac{\sqrt{2/3}}{\sqrt{36}} = \frac{\sqrt{2/3}}{6}$$

$$z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

Using the central limit theorem:

$$P(2.1 < \overline{X} < 2.5) = P\left(\frac{2.1-2}{\frac{\sqrt{2/3}}{6}} < Z < \frac{2.5-2}{\frac{\sqrt{2/3}}{6}}\right) = P(0.7348 < Z < 3.6742)$$

$$= P(Z < 3.6742) - P(Z < 0.7348) = 1 - 0.7688 = 0.2312$$

7-12.

$\mu_X = 8.2$ minutes	$n = 49$
$\sigma_X = 1.5$ minutes	$\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{1.5}{\sqrt{49}} = 0.21$ $\mu_{\overline{X}} = \mu_X = 8.2$ mins

Using the central limit theorem, $\overline{X}$ is approximately normally distributed.

a) $P(\overline{X} < 10) = P(Z < \frac{10 - 8.2}{0.2143}) = P(Z < 8.4) = 1$

b) $P(5 < \overline{X} < 10) = P(\frac{5-8.2}{0.2143} < Z < \frac{10-8.2}{0.2143})$

$= P(Z < 8.4) - P(Z < -14.932) = 1 - 0 = 1$

c) $P(\overline{X} < 6) = P(Z < \frac{6-8.2}{0.2143}) = P(Z < -10.27) = 0$

7-13.

$n_1 = 16$	$n_2 = 9$	
$\mu_1 = 75$	$\mu_2 = 70$	$\overline{X}_1 - \overline{X}_2 \sim N(\mu_{\overline{X}_1} - \mu_{\overline{X}_2}, \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2}) \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$
$\sigma_1 = 8$	$\sigma_2 = 12$	
		$\sim N(75-70, \frac{8^2}{16} + \frac{12^2}{9}) \sim N(5,20)$

a) $P(\overline{X}_1 - \overline{X}_2 > 4)$

$$P(Z > \tfrac{4-5}{\sqrt{20}}) = P(Z > -0.2236) = 1 - P(Z \le -0.2236)$$

$$= 1 - 0.4115 = 0.5885$$

b) $P(3.5 \le \overline{X}_1 - \overline{X}_2 \le 5.5)$

$$P(\tfrac{3.5-5}{\sqrt{20}} \le Z \le \tfrac{5.5-5}{\sqrt{20}}) = P(Z \le 0.1118) - P(Z \le -0.3354)$$

$$= 0.5445 - 0.3687 = 0.1759$$

7-14. If $\mu_B = \mu_A$, then $\overline{X}_B - \overline{X}_A$ is approximately normal with mean 0 and variance $\frac{\sigma_B^2}{25} + \frac{\sigma_A^2}{25} = 20.48$.

Then, $P(\overline{X}_B - \overline{X}_A > 3.5) = P(Z > \frac{3.5-0}{\sqrt{20.48}}) = P(Z > 0.773) = 0.2196$

The probability that $\overline{X}_B$ exceeds $\overline{X}_A$ by 3.5 or more is not that unusual when μ_B *and* μ_A are equal. Therefore, there is not strong evidence that μ_B is greater than μ_A.

7-15. Assume approximate normal distributions.

$(\overline{X}_{high} - \overline{X}_{low}) \sim N(60-55, \frac{4^2}{16} + \frac{4^2}{16})$

$\sim N(5,2)$

$P(\overline{X}_{high} - \overline{X}_{low} \ge 2) = P(Z \ge \frac{2-5}{\sqrt{2}}) = 1 - P(Z \le -2.12) = 1 - 0.0170 = 0.983$

7-16. a) $SE_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{1.60}{\sqrt{10}} = 0.51$

b) Here $\overline{X} \sim N(7.48, 0.51)$ and the standard normal distribution is used.

$$P(6.49 < \overline{X} < 8.47) = P\left(\frac{6.49 - 7.48}{0.51} < \frac{\overline{X} - 7.48}{0.51} < \frac{8.47 - 7.48}{0.51}\right)$$

$$= P(-1.96 < Z < 1.96) = 0.975 - 0.025 = 0.95$$

c) We assume that $\overline{X}$ is normal distributed, or the sample size is sufficiently large that the central limit theorem applies.

7-17. The proportion of samples with pH below 5.0 . Proportion $= \frac{\text{Number of samples pH} < 5}{\text{Number of total samples}} = \frac{26}{39} = 0.67$

7-18. a) $\overline{X} = 19.86, S_X = 23.65$, When $n = 8, SE_{\overline{X}} = \frac{s_X}{\sqrt{n}} = \frac{23.65}{\sqrt{8}} = 8.36$

b) Using the central limit theorem, $\overline{X} \sim N(19.86, 8.36)$

$P\big((19.86 - 8.36) < \overline{X} < (19.86 + 8.36)\big) = P(-1 < Z < 1) = 0.841 - 0.159 = 0.68$

where Z is a standard normal random variable.

c) The central limit theorem applies when the sample size n is large. Here n = 8 may be too small because the distribution of the counts of maple trees is quite skewed.

7-19. a) Point estimate of the mean proton flux is $\overline{X} = 4958$

b) Point estimate of the standard deviation is $S_X = 3420$

c) Estimate of the standard error is $S_{\overline{X}} = \frac{S_X}{\sqrt{25}} = \frac{3420}{5} = 684$

d) Point estimate of the median is 3360

e) Point estimate of the proportion of readings below 5000 is 16/25

2310	2320	2010	10800	2190	3360	5640	2540	3360
11800	2010	3430	10600	7370	2160	3200	2020	2850
3500	10200	8550	9500	2260	7730	2250		

7-20. a) Let $\overline{X}$ denotes the mean miles and $E(\overline{X}) = \mu$, we further let Y denotes the additional miles

$E(\overline{X} + Y) = E(\overline{X}) + E(Y) = \mu + 5 \cdot P(\text{head}) + (-5) \cdot P(\text{tail})$

$= \mu + 5 \times 0.5 - 5 \times 0.5 = \mu + 0 = \mu$

b) Let Y denote the additional miles. The variance of Y is

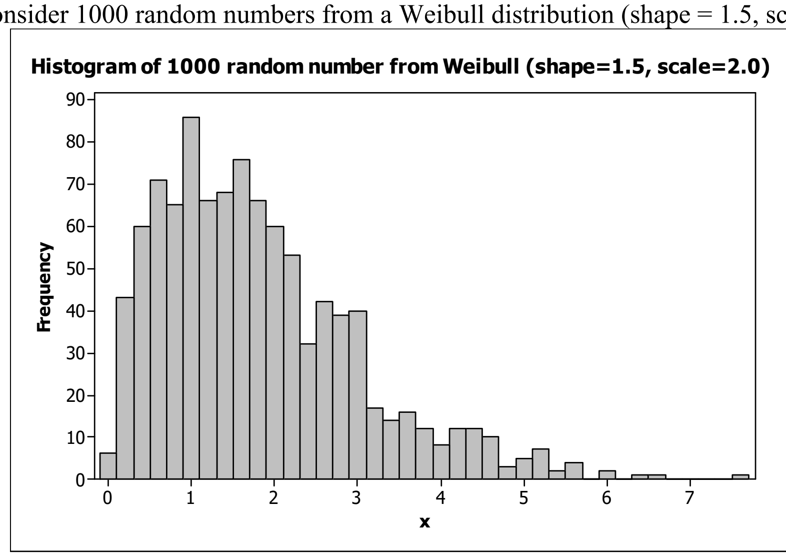
$\sigma_Y^2 = E(Y^2) - (EY)^2 = 5^2 \cdot P(\text{head}) + (-5)^2 \cdot P(\text{tail}) = \frac{25}{2} + \frac{25}{2} = 25$

$\sigma_{\overline{X}+Y}^2 = \sigma_{\overline{X}}^2 + \sigma_Y^2 = \sigma_{\overline{X}}^2 + 25$. The standard deviation of Wayne's estimator is

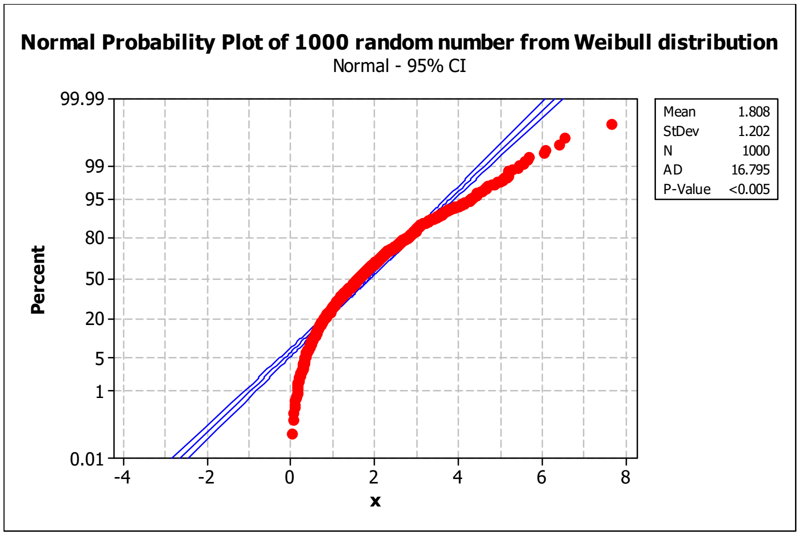
$\sqrt{\sigma_{\overline{X}}^2 + 25} > \sigma_{\overline{X}}$ and this is greater than the standard deviation of the sample mean.

c) Although Wayne's estimate is unbiased, it does not make good sense because it has a larger variance.

7-21. Consider 1000 random numbers from a Weibull distribution (shape = 1.5, scale = 2)



Normal probability plot

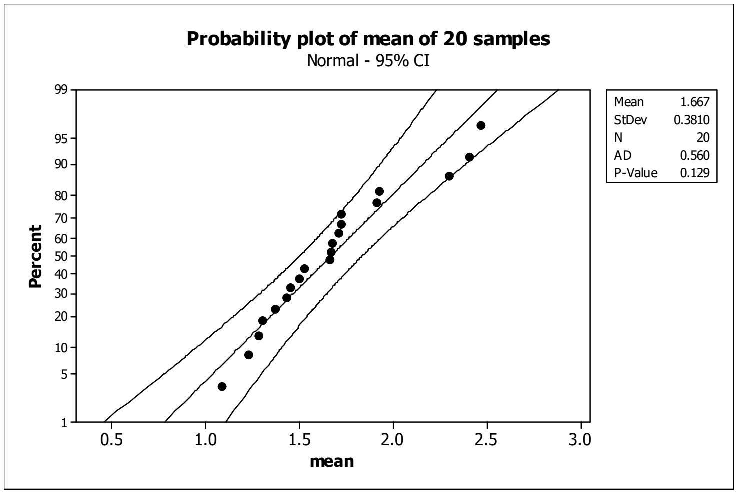


The data do not appear normally distributed. This Weibull distribution is skewed and has a long upper tail. Construct a table similar to Table 7-1

Obs	1	2	3	4	5	6	7	8	9	10
1	0.59	2.29	0.73	0.52	0.58	0.52	2.09	1.15	3.13	0.44
2	1.62	2.59	1.63	0.14	2.53	0.16	0.95	4.06	1.91	0.06
3	1.15	3.67	1.71	1.99	0.52	2.40	1.53	1.72	0.27	1.17
4	0.54	3.23	2.03	1.18	0.71	0.59	1.04	1.64	1.09	1.19
5	1.24	3.94	0.51	0.47	0.24	1.22	0.11	1.09	0.87	0.60
6	1.09	1.39	3.13	0.82	1.79	3.32	0.76	1.87	2.39	1.98
7	3.63	0.34	2.33	4.20	4.03	1.57	0.60	1.00	1.70	1.71
8	1.12	2.81	0.43	0.30	0.15	2.88	0.22	1.03	1.70	0.86
9	1.11	0.62	2.70	1.56	1.90	2.84	4.11	1.44	0.74	3.68
10	2.88	2.09	1.47	1.85	0.39	1.57	2.91	1.75	1.46	0.60
mean	1.50	2.30	1.67	1.30	1.28	1.71	1.43	1.68	1.53	1.23

Obs	11	12	13	14	15	16	17	18	19	20
1	3.19	1.27	2.03	0.97	1.92	3.28	1.68	1.75	3.08	0.10
2	0.82	1.12	0.39	3.78	3.69	1.22	1.45	1.39	3.97	0.80
3	3.55	0.78	0.79	2.05	3.89	1.51	0.34	3.57	2.03	0.66
4	3.36	0.06	1.80	0.62	1.18	0.73	0.67	0.42	1.99	3.04
5	0.47	4.51	3.61	2.02	1.71	0.71	4.08	3.76	2.77	0.12
6	0.27	1.23	0.94	3.68	3.08	0.70	0.09	1.78	2.53	2.72
7	1.85	1.89	1.77	0.38	3.05	2.35	1.00	2.27	1.34	0.48
8	0.70	1.93	1.16	0.40	3.07	1.24	3.86	0.78	3.04	0.71
9	4.16	2.10	0.20	2.98	0.48	1.66	2.44	1.51	0.57	0.69
10	0.89	1.73	1.01	0.36	2.57	1.15	1.64	1.91	2.71	1.54
mean	1.93	1.66	1.37	1.72	2.46	1.45	1.73	1.91	2.40	1.09

The normal probability plot of sample mean from each sample is much more normally distributed than the raw data.



Section 7-3

7-22. a) SE Mean $= \sigma_{\overline{X}} = \frac{S}{\sqrt{n}} = \frac{1.816}{\sqrt{20}} = 0.406$, Variance $= \sigma^2 = 1.816^2 = 3.298$

b) Estimate of mean of population = sample mean = 50.184

7-23. a) $\frac{S}{\sqrt{N}} = \text{SE Mean} \rightarrow \frac{10.25}{\sqrt{N}} = 2.05 \rightarrow N = 25$

Mean $= \frac{3761.70}{25} = 150.468$, Variance $= S^2 = 10.25^2 = 105.0625$

Variance $= \frac{\text{Sum of Squares}}{n-1} \rightarrow 105.0625 = \frac{SS}{25-1} \rightarrow SS = 2521.5$

b) Estimate of population mean = sample mean = 150.468

7-24. a) $E(\hat{\Theta}_1) = \mathrm{E}\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{2}[\mathrm{E}(X_1) + \mathrm{E}(X_2)] = \frac{1}{2}[\mu + \mu] = \mu$

Therefore, $\hat{\Theta}_1$ is an unbiased estimator of μ

$E(\hat{\Theta}_2) = \mathrm{E}\left(\frac{X_1 + 3X_2}{4}\right) = \frac{1}{4}[\mathrm{E}(X_1) + 3\mathrm{E}(X_2)] = \frac{1}{4}[\mu + 3\mu] = \mu$

Therefore $\hat{\Theta}_2$ is an unbiased estimator of μ

b) $V(\hat{\Theta}_1) = \mathrm{V}\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{4}[\mathrm{V}(X_1) + \mathrm{V}(X_2)] = \frac{1}{4}[\sigma^2 + \sigma^2] = \frac{\sigma^2}{2}$

$V(\hat{\Theta}_2) = \mathrm{V}\left(\frac{X_1 + 3X_2}{4}\right) = \frac{1}{16}[\mathrm{V}(X_1) + 3^2\mathrm{V}(X_2)] = \frac{1}{16}[\sigma^2 + 9\sigma^2] = \frac{5\sigma^2}{8}$

7-25. $E(\hat{\Theta}) = E\left(\sum_{i=1}^{n}(X_i - \overline{X})^2/c\right) = (n-1)\sigma^2/c$

Bias $= E(\hat{\Theta}) - \theta = \frac{(n-1)\sigma^2}{c} - \sigma^2 = \sigma^2(\frac{n-1}{c} - 1)$

7-26. $$\mathrm{E}(\overline{X}_1) = E\left(\frac{\sum_{i=1}^{2n} X_i}{2n}\right) = \frac{1}{2n}E\left(\sum_{i=1}^{2n} X_i\right) = \frac{1}{2n}(2n\mu) = \mu$$

$$E\left(\overline{X}_2\right) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}(n\mu) = \mu$$

$\overline{X}_1$ and $\overline{X}_2$ are unbiased estimators of μ.

The variances are $V\left(\overline{X}_1\right) = \dfrac{\sigma^2}{2n}$ and $V\left(\overline{X}_2\right) = \dfrac{\sigma^2}{n}$; compare the MSE (variance in this case),

$$\frac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)} = \frac{\sigma^2/2n}{\sigma^2/n} = \frac{n}{2n} = \frac{1}{2}$$

Because both estimators are unbiased, one concludes that $\overline{X}_1$ is the "better" estimator with the smaller variance.

7-27. $E\left(\hat{\Theta}_1\right) = \frac{1}{7}\left[E(X_1) + E(X_2) + \cdots + E(X_7)\right] = \frac{1}{7}(7E(X)) = \frac{1}{7}(7\mu) = \mu$

$$E\left(\hat{\Theta}_2\right) = \frac{1}{2}\left[E(2X_1) + E(X_6) + E(X_7)\right] = \frac{1}{2}[2\mu - \mu + \mu] = \mu$$

a) Both $\hat{\Theta}_1$ and $\hat{\Theta}_2$ are unbiased estimates of μ because the expected values of these statistics are equivalent to the true mean, μ.

b)

$$V\left(\hat{\Theta}_1\right) = V\left[\frac{X_1 + X_2 + ... + X_7}{7}\right] = \frac{1}{7^2}\left(V(X_1) + V(X_2) + \cdots + V(X_7)\right)$$

$$= \frac{1}{49}(7\sigma^2) = \frac{1}{7}\sigma^2$$

$$V(\hat{\Theta}_1) = \frac{\sigma^2}{7}$$

$$V\left(\hat{\Theta}_2\right) = V\left[\frac{2X_1 - X_6 + X_4}{2}\right] = \frac{1}{2^2}\left(V(2X_1) + V(X_6) + V(X_4)\right)$$

$$= \frac{1}{4}(4V(X_1) + V(X_6) + V(X_4))$$

$$= \frac{1}{4}\left(4\sigma^2 + \sigma^2 + \sigma^2\right) = \frac{1}{4}(6\sigma^2)$$

$$V(\hat{\Theta}_2) = \frac{3\sigma^2}{2}$$

Because both estimators are unbiased, the variances can be compared to select the better estimator. Because the variance of $\hat{\Theta}_1$ is smaller than that of $\hat{\Theta}_2$, $\hat{\Theta}_1$ is the better estimator.

7-28. Because both $\hat{\Theta}_1$ and $\hat{\Theta}_2$ are unbiased, the variances of the estimators can compared to select the better estimator. Because the variance of $\hat{\theta}_2$ is smaller than that of $\hat{\theta}_1$, $\hat{\theta}_2$ is the better estimator.

Relative Efficiency = $\dfrac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)} = \dfrac{V(\bar{\Theta}_1)}{V(\hat{\Theta}_2)} = \dfrac{10}{4} = 2.5$

7-29. $E(\hat{\Theta}_1) = \theta \qquad E(\hat{\Theta}_2) = \theta / 2$

$Bias = E(\hat{\Theta}_2) - \theta$

$= \dfrac{\theta}{2} - \theta = -\dfrac{\theta}{2}$

$V(\hat{\Theta}_1) = 10 \qquad V(\hat{\Theta}_2) = 4$

For unbiasedness, use $\hat{\Theta}_1$ because it is the only unbiased estimator.

As for minimum variance and efficiency we have

Relative Efficiency $= \dfrac{(V(\hat{\Theta}_1) + Bias^2)_1}{(V(\hat{\Theta}_2) + Bias^2)_2}$ where bias for θ_1 is 0.

Thus,

Relative Efficiency $= \dfrac{(10+0)}{\left(4 + \left(\dfrac{-\theta}{2}\right)^2\right)} = \dfrac{40}{\left(16+\theta^2\right)}$

If the relative efficiency is less than or equal to 1, $\hat{\Theta}_1$ is the better estimator.

Use $\hat{\Theta}_1$, when $\dfrac{40}{(16+\theta^2)} \le 1$

$$40 \le (16+\theta^2)$$
$$24 \le \theta^2$$
$$\theta \le -4.899 \text{ or } \theta \ge 4.899$$

If $-4.899 < \theta < 4.899$ then use $\hat{\Theta}_2$.

For unbiasedness, use $\hat{\Theta}_1$. For efficiency, use $\hat{\Theta}_1$ when $\theta \le -4.899$ or $\theta \ge 4.899$ and use $\hat{\Theta}_2$ when $-4.899 < \theta < 4.899$.

7-30. $E(\hat{\Theta}_1) = \theta$ No bias $V(\hat{\Theta}_1) = 12 = MSE(\hat{\Theta}_1)$

$E(\hat{\Theta}_2) = \theta$ No bias $V(\hat{\Theta}_2) = 10 = MSE(\hat{\Theta}_2)$

$E(\hat{\Theta}_3) \ne \theta$ Bias $MSE(\hat{\Theta}_3) = 6$ [note that this includes (bias^2)]

To compare the three estimators, calculate the relative efficiencies:

$\dfrac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)} = \dfrac{12}{10} = 1.2$, because rel. eff. > 1 use $\hat{\Theta}_2$ as the estimator for θ

$\dfrac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_3)} = \dfrac{12}{6} = 2$, because rel. eff. > 1 use $\hat{\Theta}_3$ as the estimator for θ

$\dfrac{MSE(\hat{\Theta}_2)}{MSE(\hat{\Theta}_3)} = \dfrac{10}{6} = 1.8$, because rel. eff. > 1 use $\hat{\Theta}_3$ as the estimator for θ

Conclusion: $\hat{\Theta}_3$ is the most efficient estimator, but it is biased. $\hat{\Theta}_2$ is the best "unbiased" estimator.

7-31. $n_1 = 20$, $n_2 = 10$, $n_3 = 8$

Show that S^2 is unbiased.

$$E(S^2) = E\left(\frac{20S_1^2 + 10S_2^2 + 8S_3^2}{38}\right)$$
$$= \frac{1}{38}\left(E(20S_1^2) + E(10S_2^2) + E(8S_3^2)\right)$$
$$= \frac{1}{38}\left(20\sigma_1^2 + 10\sigma_2^2 + 8\sigma_3^2\right) = \frac{1}{38}\left(38\sigma^2\right) = \sigma^2$$

Therefore, S^2 is an unbiased estimator of σ^2 .

7-32. Show that $\dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n}$ is a biased estimator of σ^2

a)
$$E\left(\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n}\right)$$
$$= \frac{1}{n}E\left(\sum_{i=1}^{n}\left(X_i - n\overline{X}\right)^2\right) = \frac{1}{n}\left(\sum_{i=1}^{n}E\left(X_i^2\right) - nE\left(\overline{X}^2\right)\right) = \frac{1}{n}\left(\sum_{i=1}^{n}\left(\mu^2 + \sigma^2\right) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right)$$
$$= \frac{1}{n}\left(n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2\right) = \frac{1}{n}((n-1)\sigma^2) = \sigma^2 - \frac{\sigma^2}{n}$$

Therefore $\dfrac{\sum\left(X_i - \overline{X}\right)^2}{n}$ is a biased estimator of σ^2

b) Bias $= E\left[\dfrac{\sum\left(X_i^2 - n\overline{X}\right)^2}{n}\right] - \sigma^2 = \sigma^2 - \dfrac{\sigma^2}{n} - \sigma^2 = -\dfrac{\sigma^2}{n}$

c) Bias decreases as n increases.

7-33. a) Show that $\overline{X}^2$ is a biased estimator of μ^2. Using $\mathrm{E}\left(\mathrm{X}^2\right) = \mathrm{V(X)} + \left[\mathrm{E(X)}\right]^2$

$$E\left(\overline{X}^2\right) = \frac{1}{n^2}E\left(\sum_{i=1}^{n}X_i\right)^2 = \frac{1}{n^2}\left(V\left(\sum_{i=1}^{n}X_i\right) + \left[E\left(\sum_{i=1}^{n}X_i\right)\right]^2\right)$$
$$= \frac{1}{n^2}\left(n\sigma^2 + \left(\sum_{i=1}^{n}\mu\right)^2\right) = \frac{1}{n^2}\left(n\sigma^2 + (n\mu)^2\right)$$
$$= \frac{1}{n^2}\left(n\sigma^2 + n^2\mu^2\right)E\left(\overline{X}^2\right) = \frac{\sigma^2}{n} + \mu^2$$

Therefore, $\overline{X}^2$ is a biased estimator of μ.^{2}

b) Bias = $E\left(\overline{X}^2\right) - \mu^2 = \dfrac{\sigma^2}{n} + \mu^2 - \mu^2 = \dfrac{\sigma^2}{n}$

c) Bias decreases as *n* increases.

7-34. a) The average of the 26 observations provided can be used as an estimator of the mean pull force because we know it is unbiased. This value is 75.615 pounds.

b) The median of the sample can be used as an estimate of the point that divides the population into a "weak" and "strong" half. This estimate is 75.2 pounds.

c) Our estimate of the population variance is the sample variance or 2.738 square pounds. Similarly, our estimate of the population standard deviation is the sample standard deviation or 1.655 pounds.

d) The estimated standard error of the mean pull force is $1.655/26^{1/2} = 0.325$. This value is the standard deviation, not of the pull force, but of the mean pull force of the sample.

e) Only one connector in the sample has a pull force measurement under 73 pounds. Our point estimate for the proportion requested is then 1/26 = 0.0385

7-35. **Descriptive Statistics**

```
Variable          N   Mean      Median     TrMean      StDev    SE Mean
Oxide Thickness 24  423.33     424.00     423.36       9.08       1.85
```

a) The mean oxide thickness, as estimated by Minitab from the sample, is 423.33 Angstroms.
b) The standard deviation for the population can be estimated by the sample standard deviation, or 9.08 Angstroms.
c) The standard error of the mean is 1.85 Angstroms.
d) Our estimate for the median is 424 Angstroms.
e) Seven of the measurements exceed 430 Angstroms, so our estimate of the proportion requested is 7/24 = 0.2917

425	431	416	419	421	436	418	410					
431	433	423	426	410	435	436	428	411	426	409	437	
422	428	413	416									

7-36. a) $E(\hat{p}) = E(X/n) = \frac{1}{n}E(X) = \frac{1}{n}np = p$

b) The variance of $\hat{p}$ is $\frac{p(1-p)}{n}$ so its standard error must be $\sqrt{\frac{p(1-p)}{n}}$. To estimate this parameter we substitute our estimate of p into it.

7-37. a) $E(\overline{X}_1 - \overline{X}_2) = E(\overline{X}_1) - E(\overline{X}_2) = \mu_1 - \mu_2$

b) $s.e. = \sqrt{V(\overline{X}_1 - \overline{X}_2)} = \sqrt{V(\overline{X}_1) + V(\overline{X}_2) + 2COV(\overline{X}_1, \overline{X}_2)} = \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}$

This standard error can be estimated by using the estimates for the standard deviations of populations 1 and 2.

c)

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

$$= \frac{1}{n_1+n_2-2}\left[(n_1-1)\cdot\sigma_1^{\ 2} + (n_2-1)\cdot\sigma_2^{\ 2}\right] = \frac{n_1+n_2-2}{n_1+n_2-2}\sigma^2 = \sigma^2$$

7-38. a) $E(\hat{\mu}) = E(\alpha\overline{X}_1 + (1-\alpha)\overline{X}_2) = \alpha E(\overline{X}_1) + (1-\alpha)E(\overline{X}_2) = \alpha\mu + (1-\alpha)\mu = \mu$

b)

$$s.e.(\hat{\mu}) = \sqrt{V(\alpha\overline{X}_1 + (1-\alpha)\overline{X}_2)} = \sqrt{\alpha^2 V(\overline{X}_1) + (1-\alpha)^2 V(\overline{X}_2)}$$
$$= \sqrt{\alpha^2 \frac{\sigma_1{}^2}{n_1} + (1-\alpha)^2 \frac{\sigma_2{}^2}{n_2}} = \sqrt{\alpha^2 \frac{\sigma_1{}^2}{n_1} + (1-\alpha)^2 a \frac{\sigma_1{}^2}{n_2}}$$
$$= \sigma_1 \sqrt{\frac{\alpha^2 n_2 + (1-\alpha)^2 a n_1}{n_1 n_2}}$$

c) The value of alpha that minimizes the standard error is $\alpha = \dfrac{an_1}{n_2 + an_1}$

d) With $a = 4$ and $n_1=2n_2$, the value of α to choose is 8/9. The arbitrary value of α = 0.5 is too small and results in a larger standard error. With α = 8/9, the standard error is

$$s.e.(\hat{\mu}) = \sigma_1 \sqrt{\frac{(8/9)^2 n_2 + (1/9)^2 8n_2}{2n_2{}^2}} = \frac{0.667\sigma_1}{\sqrt{n_2}}$$

If α = 0.5 the standard error is

$$s.e.(\hat{\mu}) = \sigma_1 \sqrt{\frac{(0.5)^2 n_2 + (0.5)^2 8n_2}{2n_2{}^2}} = \frac{1.0607\sigma_1}{\sqrt{n_2}}$$

7-39.

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

b) $\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

c) An estimate of the standard error could be obtained substituting $\frac{X_1}{n_1}$ for p_1 and $\frac{X_2}{n_2}$ for p_2 in the equation shown in (b).

d) Our estimate of the difference in proportions is 0.01

e) The estimated standard error is 0.0413

7-40. $X \sim lognormal(1.5, 0.8^2)$, $n = 15, n_B = 200$, The original samples (n = 15):

#	1	2	3	4	5
Value	15.76	1.47	4.94	2.16	8.88
#	6	7	8	9	10
Value	1.32	13.40	4.14	4.10	4.19
#	11	12	13	14	15
Value	1.97	0.69	2.87	2.01	5.52

Here 200 bootstrap samples are generated and the sample median is computed for each. The standard error is estimated as the standard deviation of these sample medians, and in this case we obtain 1.04.

R code:

```
#generate the first 15 samples from the target distribution
n=15;
sample0=rlnorm(n, meanlog = 1.5, sdlog = 0.8);
#generate bootstrap samples and medians
nb=200
```

```
data=matrix(0, nb,n);
med=matrix(0, nb,1);
for (i in c(1: nb))
{
data[i,]=sample(sample0, n, replace = TRUE, prob = NULL)
med[i]=median(data[i,])
}
#calculate the standard error of the sample median "se"
se = sd(med)
```

7-41. $X \sim exp(\lambda = 0.1), n = 8, n_B = 100$, the original sample (n = 8):

#	1	2	3	4	5	6	7	8
Value	1.88	4.27	18.15	8.37	26.25	5.76	0.74	7.45

Here 100 bootstrap samples are generated and the sample median is computed for each. The standard error is estimated as the standard deviation of these sample medians, and in this case we obtain 2.85.

R code:

```
#generate the first 8 samples from the target distribution
n=8
sample0=rexp(n, rate = 0.1)
#generate bootstrap samples and medians
nb=100
data=matrix(0,nb,n)
med=matrix(0,nb,1);
for (i in c(1:nb))
{
data[i,]=sample(sample0, n, replace = TRUE, prob = NULL)
med[i]=median(data[i,])
}
#calculate the standard error of the sample median "se"
se=sd(med)
```

7-42. $X \sim norm(\mu = 10,\ \sigma^2 = 4^2), n = 16,\ n_B = 200$, the original sample (n = 16):

#	1	2	3	4	5	6	7	8
Value	4.26	6.59	12.36	7.47	10.84	1.17	17.02	14.10
#	9	10	11	12	13	14	15	16
Value	12.30	6.73	16.75	11.87	12.25	7.52	7.10	9.80

Here 200 bootstrap samples are generated and the sample mean is computed for each. The standard error is estimated as the standard deviation of these sample medians, and in this case we obtain 0.96. The bootstrap result is near 1, the true standard error.

R code

```
#generate the first 8 samples from the target distribution
n=16
nb=200
sample0=rnorm(n, mean = 10, sd=4)
#generate bootstrap samples and means
data=matrix(0,nb,n)
ave=matrix(0,nb,1);
```

```
for (i in c(1:nb))
{
data[i,]=sample(sample0, n, replace = TRUE, prob = NULL)
ave[i]=mean(data[i,])
}
#calculate the standard error of the sample mean
se=sd(ave)
```

7-43. Suppose that two independent random samples (of size n_1 and n_2) from two normal distributions are available. Explain how you would estimate the standard error of the difference in sample means $\overline{X}_1 - \overline{X}_2$ with the bootstrap method.

One case use the fact that $V(\overline{X}_1 - \overline{X}_2) = V(\overline{X}_1) + V(\overline{X}_2)$ (because the samples are independent) so that the standard error of $\overline{X}_1 - \overline{X}_2$ is $\sqrt{V(\overline{X}_1) + V(\overline{X}_2)}$. Then the bootstrap method can be used as in the previous exercise to estimate $V(\overline{X}_1)$ and $V(\overline{X}_2)$.

Section 7-4

7-44. $f(x) = p(1-p)^{x-1}$

$$L(p) = \prod_{i=1}^{n} p(1-p)^{x_i - 1} = p^n (1-p)^{\sum_{i=1}^{n} x_i - n}$$

$$\ln L(p) = n \ln p + \left(\sum_{i=1}^{n} x_i - n \right) \ln(1-p)$$

$$\frac{\partial \ln L(p)}{\partial p} = \frac{n}{p} - \frac{\sum_{i=1}^{n} x_i - n}{1-p} \equiv 0$$

$$0 = \frac{(1-p)n - p\left(\sum_{i=1}^{n} x_i - n \right)}{p(1-p)} = \frac{n - np - p\sum_{i=1}^{n} x_i + pn}{p(1-p)}$$

$$0 = n - p\sum_{i=1}^{n} x_i$$

$$\hat{p} = \frac{n}{\sum_{i=1}^{n} x_i}$$

7-45. $f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$ $\quad L(\lambda) = \prod_{i=1}^{n} \dfrac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \dfrac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$

$$\ln L(\lambda) = -n\lambda \ln e + \sum_{i=1}^{n} x_i \ln \lambda - \sum_{i=1}^{n} \ln x_i!$$

$$\frac{d \ln L(\lambda)}{d\lambda} = -n + \frac{1}{\lambda}\sum_{i=1}^{n} x_i \equiv 0$$

$$= -n + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0$$

$$\sum_{i=1}^{n} x_i = n\lambda$$

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n}$$

7-46. $f(x) = (\theta+1)x^{\theta}$

$$L(\theta) = \prod_{i=1}^{n} (\theta+1)x_i^{\theta} = (\theta+1)x_1^{\theta} \times (\theta+1)x_2^{\theta} \times ... = (\theta+1)^n \prod_{i=1}^{n} x_i^{\theta}$$

$$\ln L(\theta) = n\ln(\theta+1) + \theta \ln x_1 + \theta \ln x_2 + ... = n\ln(\theta+1) + \theta \sum_{i=1}^{n} \ln x_i$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{n}{\theta+1} + \sum_{i=1}^{n} \ln x_i = 0$$

$$\frac{n}{\theta+1} = -\sum_{i=1}^{n} \ln x_i$$

$$\hat{\theta} = \frac{n}{-\sum_{i=1}^{n} \ln x_i} - 1$$

7-47. $f(x) = \lambda e^{-\lambda(x-\theta)}$ for $x \ge \theta$ $\quad L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda(x-\theta)} = \lambda^n e^{-\lambda \sum_{i=1}^{n}(x-\theta)} = \lambda^n e^{-\lambda\left(\sum_{i=1}^{n} x - n\theta\right)}$

$$\ln L(\lambda, \theta) = n\ln\lambda - \lambda\sum_{i=1}^{n} x_i + \lambda n\theta$$

$$\frac{d \ln L(\lambda,\theta)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i + n\theta \equiv 0$$

$$\frac{n}{\lambda} = \sum_{i=1}^{n} x_i - n\theta$$

$$\hat{\lambda} = n \Big/ \left(\sum_{i=1}^{n} x_i - n\theta\right)$$

$$\hat{\lambda} = \frac{1}{\bar{x} - \theta}$$

The other parameter θ cannot be estimated by setting the derivative of the log likelihood with respect to θ to zero because the log likelihood is a linear function of θ. The range of the likelihood is important.

The joint density function and therefore the likelihood is zero for $\theta < Min(X_1, X_2, \ldots, X_n)$. The term in the log likelihood $-n\lambda\theta$ is maximized for θ as small as possible within the range of nonzero likelihood. Therefore, the log likelihood is maximized for θ estimated with $Min(X_1, X_2, \ldots, X_n)$ so that $\hat{\theta} = x_{\min}$

b) Example: Consider traffic flow and let the time that has elapsed between one car passing a fixed point and the instant that the next car begins to pass that point be considered time headway. This headway can be modeled by the shifted exponential distribution.
Example in Reliability: Consider a process where failures are of interest. Suppose that a unit is put into operation at $x = 0$, but no failures will occur until θ time units of operation. Failures will occur only after the time θ.

7-48.

$$L(\theta) = \prod_{i=1}^{n} \frac{x_i e^{-x_i/\theta}}{\theta^2} \qquad \ln L(\theta) = \sum \ln(x_i) - \sum \frac{x_i}{\theta} - 2n\ln\theta$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{1}{\theta^2}\sum x_i - \frac{2n}{\theta}$$

Setting the last equation equal to zero and solving for theta yields

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{2n}$$

7-49. $E(X) = \dfrac{a-0}{2} = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$, therefore: $\hat{a} = 2\overline{X}$

The expected value of this estimate is the true parameter, so it is unbiased. This estimate is reasonable in one sense because it is unbiased. However, there are obvious problems. Consider the sample $x_1=1$, $x_2 = 2$ and $x_3 =10$. Now $\overline{x} = 4.37$ and $\hat{a} = 2\overline{x} = 8.667$. This is an unreasonable estimate of a, because clearly $a \geq 10$.

7-50. a) $\int_{-1}^{1} c(1+\theta x)dx = 1 = (cx + c\theta \frac{x^2}{2})\Big|_{-1}^{1} = 2c$

so that the constant c should equal 0.5

b) $E(X) = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \dfrac{\theta}{3}$ $\qquad \hat{\theta} = 3 \cdot \dfrac{1}{n}\sum_{i=1}^{n} X_i$

c) $E(\hat{\theta}) = E\left(3 \cdot \frac{1}{n}\sum_{i=1}^{n} X_i\right) = E(3\overline{X}) = 3E(\overline{X}) = 3\frac{\theta}{3} = \theta$

d)

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{2}(1+\theta X_i) \qquad \ln L(\theta) = n\ln(\frac{1}{2}) + \sum_{i=1}^{n} \ln(1+\theta X_i)$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{n} \frac{X_i}{(1+\theta X_i)}$$

By inspection, the value of θ that maximizes the likelihood is max (X_i)

7-51. a) $E(X^2) = 2\theta = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ so $\hat{\Theta} = \frac{1}{2n}\sum_{i=1}^{n} X_i^2$

b)

$$L(\theta) = \prod_{i=1}^{n} \frac{x_i e^{-x_i^2/2\theta}}{\theta} \qquad \ln L(\theta) = \sum \ln(x_i) - \sum \frac{x_i^2}{2\theta} - n\ln\theta$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{1}{2\theta^2}\sum x_i^2 - \frac{n}{\theta}$$

Setting the last equation equal to zero, the maximum likelihood estimate is

$$\hat{\Theta} = \frac{1}{2n}\sum_{i=1}^{n} X_i^2$$

and this is the same result obtained in part (a)

c)

$$\int_0^a f(x)dx = 0.5 = 1 - e^{-a^2/2\theta}$$

$$a = \sqrt{-2\theta \ln(0.5)} = \sqrt{2\theta \ln(2)}$$

We can estimate the median (*a*) by substituting our estimate for θ into the equation for *a*.

7-52. a) $\hat{a}$ cannot be unbiased since it will always be less than a.

b) bias = $\frac{na}{n+1} - \frac{a(n+1)}{n+1} = -\frac{a}{n+1} \underset{n\to\infty}{\to} 0$.

c) $2\overline{X}$

d) P(Y≤y)=P(X_1, …,X_n ≤ y)=[P(X_1≤y)]^n=$\left(\frac{y}{a}\right)^n$. Thus, f(y) is as given. Thus,

bias = E(Y) – a = $\frac{an}{n+1} - a = -\frac{a}{n+1}$.

e) For any $n > 1$, $n(n+2) > 3n$ so the variance of $\hat{a}_2$ is less than that of $\hat{a}_1$. It is in this sense that the second estimator is better than the first.

7-53. a)

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

$$\ln L(\beta,\delta)=\sum_{i=1}^{n}\ln\left[\frac{\beta}{\delta}\left(\frac{x_i}{\delta}\right)^{\beta-1}\right]-\sum\left(\frac{x_i}{\delta}\right)^{\beta}=n\ln\left(\frac{\beta}{\delta}\right)+(\beta-1)\sum\ln\left(\frac{x_i}{\delta}\right)-\sum\left(\frac{x_i}{\delta}\right)^{\beta}$$

b)

$$\frac{\partial \ln L(\beta,\delta)}{\partial\beta}=\frac{n}{\beta}+\sum\ln\left(\frac{x_i}{\delta}\right)-\sum\ln\left(\frac{x_i}{\delta}\right)\left(\frac{x_i}{\delta}\right)^{\beta}$$

$$\frac{\partial \ln L(\beta,\delta)}{\partial\delta}=-\frac{n}{\delta}-(\beta-1)\frac{n}{\delta}+\beta\frac{\sum x_i^{\beta}}{\delta^{\beta+1}}$$

Upon setting $\frac{\partial \ln L(\beta,\delta)}{\partial\delta}$ equal to zero, we obtain

$$\delta^{\beta}n=\sum x_i^{\beta} \quad \text{and} \quad \delta=\left[\frac{\sum x_i^{\beta}}{n}\right]^{1/\beta}$$

Upon setting $\frac{\partial \ln L(\beta,\delta)}{\partial\beta}$ equal to zero and substituting for δ, we obtain

$$\frac{n}{\beta}+\sum\ln x_i-n\ln\delta=\frac{1}{\delta^{\beta}}\sum x_i^{\beta}(\ln x_i-\ln\delta)$$

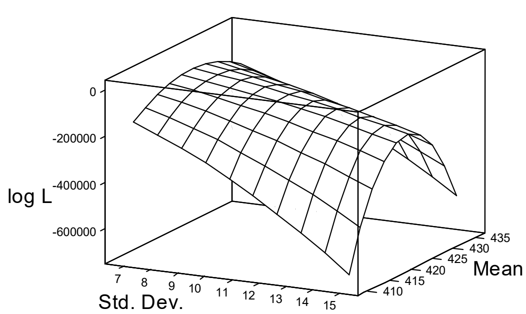
$$\frac{n}{\beta}+\sum\ln x_i-\frac{n}{\beta}\ln\left(\frac{\sum x_i^{\beta}}{n}\right)=\frac{n}{\sum x_i^{\beta}}\sum x_i^{\beta}\ln x_i-\frac{n}{\sum x_i^{\beta}}\sum x_i^{\beta}\frac{1}{\beta}\ln\left(\frac{\sum x_i^{\beta}}{n}\right)$$

and $$\frac{1}{\beta}=\left[\frac{\sum x_i^{\beta}\ln x_i}{\sum x_i^{\beta}}+\frac{\sum\ln x_i}{n}\right]$$

c) Numerical iteration is required.

7-54. a) Using the results from the example, we obtain that the estimate of the mean is 423.33 and the estimate of the variance is 82.4464

b)



The function has an approximate ridge and its curvature is not too pronounced. The maximum value for standard deviation is at 9.08, although it is difficult to see on the graph.

c) When n is increased to 40, the graph looks the same although the curvature is more pronounced. As n increases, it is easier to determine the maximum value for the standard deviation is on the graph.

7-55. From the example, the posterior distribution for μ is normal with mean $\dfrac{(\sigma^2/n)\mu_0 + \sigma_0^2\bar{x}}{\sigma_0^2 + \sigma^2/n}$ and variance $\dfrac{\sigma_0^2/(\sigma^2/n)}{\sigma_0^2 + \sigma^2/n}$. The Bayes estimator for μ goes to the MLE as n increases. This follows because σ^2/n goes to 0, and the estimator approaches $\dfrac{\sigma_0^2\bar{x}}{\sigma_0^2}$ (the σ_0^2's cancel). Thus, in the limit $\hat{\mu} = \bar{x}$.

7-56. a) Because $f(x\mid\mu) = \dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ and $f(\mu) = \dfrac{1}{b-a}$ for $a \le \mu \le b$, the joint distribution is

$$f(x,\mu) = \frac{1}{(b-a)\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ for } -\infty < x < \infty \text{ and } a \le \mu \le b.$$

Then, $f(x) = \dfrac{1}{b-a}\displaystyle\int_a^b \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,d\mu$

and this integral is recognized as a normal probability. Therefore,

$$f(x) = \frac{1}{b-a}\left[\Phi\left(\tfrac{b-x}{\sigma}\right) - \Phi\left(\tfrac{a-x}{\sigma}\right)\right]$$

where Φ(x) is the standard normal cumulative distribution function. Then

$$f(\mu\mid x) = \frac{f(x,\mu)}{f(x)} = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma\left[\Phi\left(\tfrac{b-x}{\sigma}\right) - \Phi\left(\tfrac{a-x}{\sigma}\right)\right]}$$

b) The Bayes estimator is

$$\tilde{\mu} = \int_a^b \frac{\mu e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,d\mu}{\sqrt{2\pi}\sigma\left[\Phi\left(\tfrac{b-x}{\sigma}\right) - \Phi\left(\tfrac{a-x}{\sigma}\right)\right]}.$$

Let v = (x - μ). Then, dv = - dμ and

$$\tilde{\mu} = \int_{x-b}^{x-a} \frac{(x-v)e^{-\frac{v^2}{2\sigma^2}}\,dv}{\sqrt{2\pi}\sigma\left[\Phi\left(\tfrac{b-x}{\sigma}\right) - \Phi\left(\tfrac{a-x}{\sigma}\right)\right]} = \frac{x\left[\Phi\left(\tfrac{x-a}{\sigma}\right) - \Phi\left(\tfrac{x-b}{\sigma}\right)\right]}{\left[\Phi\left(\tfrac{b-x}{\sigma}\right) - \Phi\left(\tfrac{a-x}{\sigma}\right)\right]} - \int_{x-b}^{x-a} \frac{v e^{-\frac{v^2}{2\sigma^2}}\,dv}{\sqrt{2\pi}\sigma\left[\Phi\left(\tfrac{b-x}{\sigma}\right) - \Phi\left(\tfrac{a-x}{\sigma}\right)\right]}$$

Let w = $\dfrac{v^2}{2\sigma^2}$. Then, $dw = [\tfrac{2v}{2\sigma^2}]dv = [\tfrac{v}{\sigma^2}]dv$ and

$$\tilde{\mu} = x - \int_{\frac{(x-b)^2}{2\sigma^2}}^{\frac{(x-a)^2}{2\sigma^2}} \frac{\sigma e^{-w}\,dw}{\sqrt{2\pi}\left[\Phi\left(\tfrac{b-x}{\sigma}\right) - \Phi\left(\tfrac{a-x}{\sigma}\right)\right]} = x + \frac{\sigma}{\sqrt{2\pi}}\left[\frac{e^{-\frac{(x-a)^2}{2\sigma^2}} - e^{-\frac{(x-b)^2}{2\sigma^2}}}{\Phi\left(\tfrac{b-x}{\sigma}\right) - \Phi\left(\tfrac{a-x}{\sigma}\right)}\right]$$

7-57. a) $f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$ for x = 0, 1, 2, and $f(\lambda) = \left(\dfrac{m+1}{\lambda_0}\right)^{m+1}\dfrac{\lambda^m e^{-(m+1)\frac{\lambda}{\lambda_0}}}{\Gamma(m+1)}$ for λ > 0.

Then,

$$f(x,\lambda) = \frac{(m+1)^{m+1}\lambda^{m+x}e^{-\lambda-(m+1)\frac{\lambda}{\lambda_0}}}{\lambda_0^{m+1}\Gamma(m+1)x!}.$$

This last density is recognized to be a gamma density as a function of λ. Therefore, the posterior distribution of λ is a gamma distribution with parameters m + x + 1 and 1 + $\frac{m+1}{\lambda_0}$.

b) The mean of the posterior distribution can be obtained from the results for the gamma distribution to be

$$\frac{m+x+1}{\left[1+\frac{m+1}{\lambda_0}\right]} = \lambda_0\left(\frac{m+x+1}{m+\lambda_0+1}\right)$$

7-58 a) From the example, the Bayes estimate is $\tilde{\mu} = \frac{\frac{9}{25}(4)+1(4.85)}{\frac{9}{25}+1} = 4.625$

b.) $\hat{\mu} = \bar{x} = 4.85$ The Bayes estimate appears to underestimate the mean.

7-59. a) From the example, $\tilde{\mu} = \frac{(0.01)(5.03)+(\frac{1}{25})(5.05)}{0.01+\frac{1}{25}} = 5.046$

b) $\hat{\mu} = \bar{x} = 5.05$ The Bayes estimate is very close to the MLE of the mean.

7-60. a) $f(x \mid \lambda) = \lambda e^{-\lambda x},\ x \geq 0$ and $f(\lambda) = 0.01e^{-0.01\lambda}$. Then,

$$f(x_1, x_2, \lambda) = \lambda^2 e^{-\lambda(x_1+x_2)}0.01e^{-0.01\lambda} = 0.01\lambda^2 e^{-\lambda(x_1+x_2+0.01)}.$$

As a function of λ, this is recognized as a gamma density with parameters 3 and $x_1 + x_2 + 0.01$. Therefore, the posterior mean for λ is

$$\tilde{\lambda} = \frac{3}{x_1+x_2+0.01} = \frac{3}{2\bar{x}+0.01} = 0.00133.$$

b) Using the Bayes estimate for λ, P(X<1000)= $\int_0^{1000} 0.00133e^{-.00133x}dx = 0.736$

Supplemental Exercises

7-61. $f(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} \quad for \quad x_1 > 0, x_2 > 0, ..., x_n > 0$

7-62. $f(x_1, x_2, x_3, x_4, x_5) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^5 \exp\left(-\sum_{i=1}^{5}\frac{(x_i-\mu)^2}{2\sigma^2}\right)$

7-63. $f(x_1, x_2, x_3, x_4) = 1 \quad for \quad 0 \le x_1 \le 1, 0 \le x_2 \le 1, 0 \le x_3 \le 1, 0 \le x_4 \le 1$

7-64. $$\overline{X}_1 - \overline{X}_2 \sim N(100-105, \frac{1.5^2}{25} + \frac{2^2}{30})$$
$$\sim N(-5, 0.2233)$$

7-65. $X \sim N(50,144)$

$$P(47 \le \overline{X} \le 53) = P\left(\frac{47-50}{12/\sqrt{36}} \le Z \le \frac{53-50}{12/\sqrt{36}}\right) = P(-1.5 \le Z \le 1.5)$$
$$= P(Z \le 1.5) - P(Z \le -1.5) = 0.9332 - 0.0668 = 0.8664$$

No, because Central Limit Theorem states that with large samples ($n \ge 30$), $\overline{X}$ is approximately normally distributed.

7-66. Assume $\overline{X}$ is approximately normally distributed.

$$P(\overline{X} > 4985) = 1 - P(\overline{X} \le 4985) = 1 - P(Z \le \frac{4985-5500}{100/\sqrt{9}})$$
$$= 1 - P(Z \le -15.45) = 1 - 0 = 1$$

7-67. $$z = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{52-50}{\sqrt{2/16}} = 5.6569$$

$P(Z > z) \approx 0$. The results are *very unusual*.

7-68. $P(\overline{X} \le 37) = P(Z \le -5.36) \approx 0$

7-69. Binomial with p equal to the proportion of defective chips and $n = 100$.

7-70. $E(a\overline{X}_1 + (1-a)\overline{X}_2 = a\mu + (1-a)\mu = \mu$

$$V(\overline{X}) = V[a\overline{X}_1 + (1-a)\overline{X}_2]$$
$$= a^2 V(\overline{X}_1) + (1-a)^2 V(\overline{X}_2) = a^2(\tfrac{\sigma^2}{n_1}) + (1 - 2a + a^2)(\tfrac{\sigma^2}{n_2})$$
$$= \frac{a^2\sigma^2}{n_1} + \frac{\sigma^2}{n_2} - \frac{2a\sigma^2}{n_2} + \frac{a^2\sigma^2}{n_2} = (n_2 a^2 + n_1 - 2n_1 a + n_1 a^2)(\frac{\sigma^2}{n_1 n_2})$$
$$\frac{\partial V(\overline{X})}{\partial a} = (\frac{\sigma^2}{n_1 n_2})(2n_2 a - 2n_1 + 2n_1 a) \equiv 0$$
$$0 = 2n_2 a - 2n_1 + 2n_1 a$$
$$2a(n_2 + n_1) = 2n_1$$
$$a(n_2 + n_1) = n_1$$
$$a = \frac{n_1}{n_2 + n_1}$$

7-71.

$$L(\theta) = \left(\frac{1}{2\theta^3}\right)^n e^{\sum_{i=1}^{n} \frac{-x_i}{\theta}} \prod_{i=1}^{n} x_i^{\,2}$$

$$\ln L(\theta) = n\ln\left(\frac{1}{2\theta^3}\right) + 2\sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\frac{x_i}{\theta}$$

$$\frac{\partial \ln L(\theta)}{\partial\theta} = \frac{-3n}{\theta} + \sum_{i=1}^{n}\frac{x_i}{\theta^2}$$

Making the last equation equal to zero and solving for θ, we obtain

$\hat{\Theta} = \dfrac{\sum_{i=1}^{n} x_i}{3n}$ as the maximum likelihood estimate.

7-72.

$$L(\theta) = \theta^n \prod_{i=1}^{n} x_i^{\theta-1}$$

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

$$\frac{\partial \ln L(\theta)}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^{n}\ln(x_i)$$

making the last equation equal to zero and solving for theta, we obtain the maximum likelihood estimate

$$\hat{\Theta} = \frac{-n}{\sum_{i=1}^{n}\ln(x_i)}$$

7-73.

$$L(\theta) = \frac{1}{\theta^n}\prod_{i=1}^{n} x_i^{\frac{1-\theta}{\theta}}$$

$$\ln L(\theta) = -n\ln\theta + \frac{1-\theta}{\theta}\sum_{i=1}^{n}\ln(x_i)$$

$$\frac{\partial \ln L(\theta)}{\partial\theta} = -\frac{n}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^{n}\ln(x_i)$$

Upon setting the last equation equal to zero and solving for the parameter of interest, we obtain the maximum likelihood estimate

$$\hat{\Theta} = -\frac{1}{n}\sum_{i=1}^{n}\ln(x_i)$$

$$E(\hat{\theta}) = E\left[-\frac{1}{n}\sum_{i=1}^{n}\ln(x_i)\right] = \frac{1}{n}E\left[-\sum_{i=1}^{n}\ln(x_i)\right] = -\frac{1}{n}\sum_{i=1}^{n}E[\ln(x_i)]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\theta = \frac{n\theta}{n} = \theta$$

$E(\ln(X_i)) = \int_0^1 (\ln x) x^{\frac{1-\theta}{\theta}} dx$ let $u = \ln x$ and $dv = x^{\frac{1-\theta}{\theta}} dx$

then, $E(\ln(X)) = -\theta \int_0^1 x^{\frac{1-\theta}{\theta}} dx = -\theta$

7-74. a) Let $E\overline{X}^2 = \theta$. Then $V(\overline{X}) = E(\overline{X}^2) - (E\overline{X})^2$. Therefore $\sigma^2/n = \theta - \mu^2$ and $\theta = \sigma^2/n + \mu^2$

Therefore, $\overline{X}^2$ is a biased estimator of the area of the square.

b) $E(\overline{X}^2 - S^2/n) = \sigma^2/n + \mu^2 - E(S^2)/n = \mu^2$

7-75. $\hat{\mu} = \bar{x} = \frac{23.1 + 15.6 + 17.4 + \ldots + 28.7}{10} = 21.86$

Demand for all 5000 houses is θ = 5000μ
$\hat{\theta} = 5000\hat{\mu} = 5000(21.86) = 109{,}300$

The proportion estimate is $\hat{p} = \frac{7}{10} = 0.7$

Mind-Expanding Exercises

7-76. $P(X_1 = 0, X_2 = 0) = \frac{M(M-1)}{N(N-1)}$

$P(X_1 = 0, X_2 = 1) = \frac{M(N-M)}{N(N-1)}$

$P(X_1 = 1, X_2 = 0) = \frac{(N-M)M}{N(N-1)}$

$P(X_1 = 1, X_2 = 1) = \frac{(N-M)(N-M-1)}{N(N-1)}$

$P(X_1 = 0) = M/N$

$P(X_1 = 1) = \frac{N-M}{N}$

$$P(X_2 = 0) = P(X_2 = 0 \mid X_1 = 0)P(X_1 = 0) + P(X_2 = 0 \mid X_1 = 1)P(X_1 = 1)$$
$$= \frac{M-1}{N-1} \times \frac{M}{N} + \frac{M}{N-1} \times \frac{N-M}{N} = \frac{M}{N}$$
$$P(X_2 = 1) = P(X_2 = 1 \mid X_1 = 0)P(X_1 = 0) + P(X_2 = 1 \mid X_1 = 1)P(X_1 = 1)$$
$$= \frac{N-M}{N-1} \times \frac{M}{N} + \frac{N-M-1}{N-1} \times \frac{N-M}{N} = \frac{N-M}{N}$$

Because $P(X_2 = 0 \mid X_1 = 0) = \frac{M-1}{N-1}$ is not equal to $P(X_2 = 0) = \frac{M}{N}$, X_1 *and* X_2 are not independent.

7-77. a)

$$c_n = \frac{\Gamma[(n-1)/2]}{\Gamma(n/2)\sqrt{2/(n-1)}}$$

b) When $n = 10$, $c_n = 1.0281$. When $n = 25$, $c_n = 1.0105$. Therefore S is a reasonably good estimator for the standard deviation even when relatively small sample sizes are used.

7-78.

a) The likelihood is

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{\frac{-(x_i-\mu_i)^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y_i-\mu_i)^2}{2\sigma^2}}$$

The log likelihood function is

$$-2\ln(L) = \sum_{i=1}^{n} [\frac{(x_i-\mu_i)^2}{\sigma^2} + \frac{(y_i-\mu_i)^2}{\sigma^2} + 4\ln(\sqrt{2\pi\sigma^2})\,]$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} [(x_i-\mu_i)^2 + (y_i-\mu_i)^2\,] + 4n\ln(\sqrt{2\pi\sigma^2})$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} [x_i^2 + y_i^2 - 2\mu_i(x_i+y_i) + 2{\mu_i}^2] + \; 4n\ln(\sqrt{2\pi}) + 2n\ln(\sigma^2)$$

Take the derivative of each μ_i and set it to zero

$$\frac{\partial\,(-2\ln(L))}{\partial\mu_i} = \frac{-2(x_i+y_i)+4\mu_i}{\sigma^2} = 0$$

to obtain

$$\hat{\mu}_i = \frac{x_i+y_i}{2}$$

To find the maximum likelihood estimator of σ^2, substitute the estimate for μ_i and take the derivative with respect to σ^2

$$\frac{\partial\,(-2\ln(L))}{\partial\sigma^2} = -\frac{1}{\sigma^4}\sum_{i=1}^{n}[(x_i-\hat{\mu}_i)^2 + (y_i-\hat{\mu}_i)^2] + \frac{2n}{\sigma^2}$$

$$\frac{\partial\,(-2\ln(L))}{\partial\sigma^2} = -\frac{1}{\sigma^4}\sum_{i=1}^{n}\frac{2(x_i-y_i)^2}{4} + \frac{2n}{\sigma^2}$$

$$= -\frac{\sum_{i=1}^{n}(x_i-y_i)^2}{2\sigma^4} + \frac{2n}{\sigma^2}$$

Set the derivative to zero and solve

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-y_i)^2}{4n}$$

b)

$$E(\hat{\sigma}^2) = \frac{1}{4n}\sum_{i=1}^{n}E(y_i-x_i)^2 = \frac{1}{4n}\sum_{i=1}^{n}E\left(y_i^2+x_i^2-2x_iy_i\right)$$

$$= \frac{1}{4n}\sum_{i=1}^{n}[E(y_i^2) + E(x_i^2) - E(2x_iy_i)] = \frac{1}{4n}\sum_{i=1}^{n}[\sigma^2+\sigma^2+0] = \frac{\sigma^2}{2}$$

Therefore, the estimator is biased. The bias is independent of *n*.

c) An unbiased estimator of σ^2 is given by $2\hat{\sigma}^2$

7-79. $P\left(|\overline{X}-\mu| \geq \frac{c\sigma}{\sqrt{n}}\right) \leq \frac{1}{c^2}$ from Chebyshev's inequality. Then, $P\left(|\overline{X}-\mu| < \frac{c\sigma}{\sqrt{n}}\right) \geq 1-\frac{1}{c^2}$. Given an ε, n and c can be chosen sufficiently large that the last probability is near 1 and $\frac{c\sigma}{\sqrt{n}}$ is equal to ε.

7-80. a) $P\left(X_{(n)} \leq t\right) = P\left(X_i \leq t \;\; for \;\; i=1,\ldots,n\right) = [F(t)]^n$

$P\left(X_{(1)} > t\right) = P\left(X_i > t \;\; for \;\; i=1,\ldots,n\right) = [1-F(t)]^n$

Then, $P\left(X_{(1)} \leq t\right) = 1-[1-F(t)]^n$

b)

$$f_{X_{(1)}}(t) = \frac{\partial}{\partial t}F_{X_{(1)}}(t) = n[1-F(t)]^{n-1}f(t)$$

$$f_{X_{(n)}}(t) = \frac{\partial}{\partial t}F_{X_{(n)}}(t) = n[F(t)]^{n-1}f(t)$$

c) $P\left(X_{(1)} = 0\right) = F_{X_{(1)}}(0) = 1-[1-F(0)]^n = 1-p^n$ because F(0) = 1 - p.

$P\left(X_{(n)} = 1\right) = 1-F_{X_{(n)}}(0) = 1-[F(0)]^n = 1-(1-p)^n$

d) $P(X \leq t) = F(t) = \Phi\left[\frac{t-\mu}{\sigma}\right]$. From a previous exercise,

$$f_{X_{(1)}}(t) = n\left\{1-\Phi\left[\frac{t-\mu}{\sigma}\right]\right\}^{n-1}\frac{1}{\sqrt{2\pi}\sigma}\;e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

$$f_{X_{(n)}}(t) = n\left\{\Phi\left[\frac{t-\mu}{\sigma}\right]\right\}^{n-1}\frac{1}{\sqrt{2\pi}\sigma}\;e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

e) $P(X \leq t) = 1-e^{-\lambda t}$

From a previous exercise,

$F_{X_{(1)}}(t) = 1-e^{-n\lambda t}$	$f_{X_{(1)}}(t) = n\lambda e^{-n\lambda t}$
$F_{X_{(n)}}(t) = [1-e^{-\lambda t}]^n$	$f_{X_{(n)}}(t) = n[1-e^{-\lambda t}]^{n-1}\lambda e^{-\lambda t}$

7-81. $P\left(F\left(X_{(n)}\right) \leq t\right) = P\left(X_{(n)} \leq F^{-1}(t)\right) = t^n$ for $0 \leq t \leq 1$ from a previous exercise.

If $Y = F(X_{(n)})$, then $f_Y(y) = ny^{n-1}, 0 \leq y \leq 1$.

Then, $E(Y) = \int_0^1 ny^n dy = \frac{n}{n+1}$

$P\left(F\left(X_{(1)}\right) \leq t\right) = P\left(X_{(1)} \leq F^{-1}(t)\right) = 1-(1-t)^n \;\; 0 \leq t \leq 1$ from a previous exercise.

If $Y = F(X_{(1)})$, then $f_Y(y) = n(1-t)^{n-1}, 0 \leq y \leq 1$.

Then, $E(Y)=\int_0^1 yn(1-y)^{n-1}dy=\frac{1}{n+1}$ where integration by parts is used. Therefore,

$$E[F(X_{(n)})]=\frac{n}{n+1} \quad and \quad E[F(X_{(1)})]=\frac{1}{n+1}$$

7-82. $$E(V)=k\sum_{i=1}^{n-1}\ [E(X_{i+1}^2)+E(X_i^2)-2E(X_iX_{i+1})]$$

$$=k\sum_{i=1}^{n-1}\ (\sigma^2+\mu^2+\sigma^2+\mu^2-2\mu^2)=k(n-1)2\sigma^2$$

Therefore, $k=\frac{1}{2(n-1)}$

7-83. a) The traditional estimate of the standard deviation, *S*, is 3.26. The mean of the sample is 13.43 so the values of $\left|X_i-\overline{X}\right|$ corresponding to the given observations are 3.43, 1.43, 4.43, 0.57, 4.57, 1.57 and 2.57. The median of these new quantities is 2.57 so the new estimate of the standard deviation is 3.81 and this value is slightly larger than the value obtained from the traditional estimator.

b) Making the first observation in the original sample equal to 50 produces the following results. The traditional estimator, *S*, is equal to 13.91. The new estimator remains unchanged.

7-84. a)

$$\begin{aligned}T_r = X_1 &+ \\ &X_1+X_2-X_1+ \\ &X_1+X_2-X_1+X_3-X_2+ \\ &\ldots+ \\ &X_1+X_2-X_1+X_3-X_2+\ldots+X_r-X_{r-1}+ \\ &(n-r)(X_1+X_2-X_1+X_3-X_2+\ldots+X_r-X_{r-1})\end{aligned}$$

Because X_1 is the minimum lifetime of *n* items, $E(X_1)=\frac{1}{n\lambda}$.

Then, X_2-X_1 is the minimum lifetime of (*n*-1) items from the memoryless property of the exponential and $E(X_2-X_1)=\frac{1}{(n-1)\lambda}$.

Similarly, $E(X_k-X_{k-1})=\frac{1}{(n-k+1)\lambda}$. Then,

$$E(T_r)=\frac{n}{n\lambda}+\frac{n-1}{(n-1)\lambda}+\ldots+\frac{n-r+1}{(n-r+1)\lambda}=\frac{r}{\lambda} \text{ and } E\left(\frac{T_r}{r}\right)=\frac{1}{\lambda}=\mu$$

b) $V(T_r/r)=1/(\lambda^2 r)$ is related to the variance of the Erlang distribution

$V(X) = r/\lambda^2$. They are related by the value ($1/r^2$). The censored variance is ($1/r^2$) times the uncensored variance.