Solution:

1. (a) $L(\theta) = (1+\theta)\prod_i x_i^{\theta}$
   (b) $\hat{\theta} = -(1 + 1/\sum_i \log x_i)$
   (c) $Var(\hat{\theta}) \geq (1+\theta)^2/n$
   (d) Note that the joint pdf of $X_1, \ldots, X_n$ can be re-written as $(1+\theta)\exp\{\theta \sum_i \log x_i\}$. Therefore, by the Factorisation Thm, $\sum_i \log x_i$ is a sufficient statistic for θ.
   (e) The ML estimate is $\hat{\theta} = -0.849$ and a 95% CI is $-0.849 \pm 1.96(1 - 0.849)/\sqrt{8}$, or $(-0.95, -0.74)$.
   (f) $\tilde{\theta} = (1 - 2\overline{X})/(\overline{X} - 1)$.

2. (a) $E(X) = 1/\lambda \in [2\sum_i x_i/b, 2\sum_i x_i/a]$ where a and b are, respectively, $\alpha/2$ and $(1-\alpha/2)$ quantiles for the $\chi^2(n-2)$ distribution.
   (b) $(1.26, 6.41)$
   (c) Same as (a).
   (d) The following expected frequencies are obtained using the cdf $F(x|\lambda = 1/\overline{x})$

[0, 1)	[1, 2)	[2, ∞)
$F(1) - F(0) = 0.625$	$F(2) - F(1) = 0.234$	$1 - F(2) = 0.141$

   The resulting chi-square statistic is

   $$\frac{(25-26)^2}{25} + \frac{(9.36-8)^2}{9.36} + \frac{(5.64-6)^2}{5.64} = 0.26$$

   which is smaller than the critical value 3.84 obtained as $P(\chi^2(1) < 3.84) = 0.95$. Therefore we cannot reject the null hypothesis that the underlying distribution is exponential.

3. (a) $H_0 : p = 0.9$ vs $H_1 : p < 0.9$. Under H_0, p-value $= P(X \leq 15) = 0.05$. Therefore, the rejection region is $\{x : x \leq 15\}$.
   (b) Since the observed number of successes is $x = 14$ we reject the null hypothesis and conclude that the true proportion of frames surviving the test is smaller than 0.9.
   (c) The Type II error probability is quite large: $\beta = P(X \geq 16 | p = 0.8) = 0.63$. This means that if $p = 0.8$ about 63% of all samples consisting of $n = 20$ frames would result in H_0 being incorrectly not rejected. This occurs because the sample size is small and 0.8 is close to the null value 0.9. To improve the test we should increase the sample size.

4. (a) $\theta | y \sim Beta(\alpha + y, n + \beta - y)$.
   (b) $E(\theta | y = 3) = (\alpha + y)/(\alpha + \beta + n) = 4/12 = 1/3$.

5. The complete table is

Source	df	Sum of Squares	Mean Square	F
Brand	3	75,092	25,031	1.7012
Error	16	235,408	14,713	
Total	19	310,500		

Since the observed F-statistic (1.7012) is smaller than the 0.95 quantile for the F-distribution with 3 and 16 degrees of freedom (3.23) we cannot reject the null hypothesis that the brands are different.

6. The second order statistic $y_2 = 0.85$ coincides with the first quartile. An approximate confidence interval is $[y_1, y_5]$ since

$$\sum_{j=1}^{5} \binom{10}{j} (0.25)^j (0.75)^{10-j} = 0.9239.$$

7. Let $p_1 > p_0$

$$\frac{L(p_0)}{L(p_1)} = \frac{\binom{n}{y} p_0^y (1-p_0)^{n-y}}{\binom{n}{y} p_1^y (1-p_1)^{n-y}} \leq k \tag{1}$$

Re-arranging and taking logarithms on both sides gives

$$y \ln \left\{ \frac{p_0(1-p_1)}{p_1(1-p_0)} \right\} \leq \ln k - n \ln \left\{ \frac{1-p_0}{1-p_1} \right\},$$

but since $p_0 < p_1$, $\ln[p_0(1-p_1)]/[p_1(1-p_0)] < 0$. Thus,

$$\frac{y}{n} \geq c = \frac{\ln k - n \ln[p_0(1-p_1)]/[p_1(1-p_0)]}{\ln[p_0(1-p_1)]/[p_1(1-p_0)]}$$

for each $p_1 > p_0$. If $H_0 : p = p_0$ and if n is large Y/n follows approximately $N(p_0, p_0(1-p_0)/n)$. Therefore $c \approx p_0 + z_\alpha \sqrt{p_0(1-p_0)/n}$.

8. (a) A critical region is $R = \{W : |W| > 38.46\}$; (b)

```
> x = c(5.9, 7.3, 5.2, 6.8, 7.1, 5.8, 5.9, 6.6, 6.4, 6.0)
> sign(x-7.1)
 [1] -1  1 -1 -1  0 -1 -1 -1 -1 -1
> rank(abs(x-7.1))
 [1]  7.5  2.0 10.0  3.0  1.0  9.0  7.5  4.0  5.0  6.0
> sum(rank(abs(x-7.1))*sign(x-7.1))
[1] -50
```

Since $|W| = 50 > 38.46$, we reject the null hypothesis $H_0 : m = 7.1$ and conclude that the median for arthritic individuals is different from that of non-arthritic individuals.

9. (a) Minimize

$$\sum_{i=1}^{n} (X_i - k_i\theta)^2$$

so that we solve

$$\sum_{i=1}^{n} 2k_i(X_i - k_i\theta) = 0$$

and hence

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

(b) As $E(X_i) = k_i\theta$

$$E(\hat{\theta}) = \frac{\sum_{i=1}^n k_i^2\theta}{\sum_{i=1}^n k_i^2} = \theta$$

so that $\hat{\theta}$ is unbiased.

(c)

$$\text{Var}(\hat{\theta}) = \frac{\sum_{i=1}^n k_i^2\sigma^2}{(\sum_{i=1}^n k_i^2)^2} = \frac{\sigma^2}{\sum_{i=1}^n k_i^2}$$

(d) From part (a) $\hat{\theta}$ must satisfy

$$\sum_{i=1}^n k_i(X_i - k_i\hat{\theta}) = 0$$

$$\begin{aligned}
\sum_{i=1}^n (x_i - k_i\theta)^2 &= \sum_{i=1}^n (x_i - k_i\hat{\theta} + k_i\hat{\theta} - k_i\theta)^2 \\
&= \sum_{i=1}^n (x_i - k_i\hat{\theta})^2 + \sum_{i=1}^n (k_i\hat{\theta} - k_i\theta)^2 \\
&+ 2\sum_{i=1}^n k_i(x_i - k_i\hat{\theta})(\hat{\theta} - \theta)
\end{aligned}$$

(e) Now, as $\text{Var}(X_i) = \sigma^2$, we see that $E(\sum_{i=1}^n (X_i - k_i\theta)^2) = n\sigma^2$ and from (c),

$$E\{\sum_{i=1}^n k_i^2(\hat{\theta} - \theta)^2\} = \sigma^2,$$

which yields the result.