Session 7: Solutions

Exercise 7.2.1
Model: $K \sim Bin(100, \pi)$; k (realisation of random variable K) = 40.
We will test $H_0: \pi = 1/2$ vs $H_1: \pi \neq 1/2$, using test statistic $-2llr(\pi_0)$.

$l(\pi) = k \log \pi + (n-k)\log(1-\pi)$ ignoring terms not in π. We know that $\hat{\pi} = k/n$.

$$\Rightarrow llr(\pi_0) = l(\pi_0) - l(\hat{\pi}) = 40 \log 0.5 + 60 \log 0.5 - 40 \log 0.4 - 60 \log 0.6 = -2.01$$
$$\Rightarrow -2\, llr(\pi_0) = 4.02$$

This is greater than $\chi^2_{1,0.95} = 3.84$, so we reject H_0 at the 5% level, with evidence at this significance level that the binomial parameter is less than 0.5.

Exercise 7.3.1
Model and hypotheses as in 7.2.1.
a) Test based on π:

$$H_0 \Rightarrow W = \left(\frac{M - \pi_0}{S}\right) \dot{\sim} N(0,1)$$

where $\pi_0 = \frac{1}{2}$; $M = \hat{\pi} = p = \frac{k}{n} = 0.4$; $S = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.4 \times 0.6}{100}} = 0.049$.
Then

$$W = \left(\frac{0.4 - 0.5}{0.049}\right) = -2.04$$

This is greater in magnitude than $Z_{0.975} = 1.96$, so we reject H_0 at the 5% level, with some evidence, as before, that $\pi < 0.5$.
b) Test based on $\beta = \log\left(\frac{\pi}{1-\pi}\right) \Rightarrow \hat{\beta} = \log\left(\frac{0.4}{0.6}\right) = \log\left(\frac{2}{3}\right)$

From lecture notes, $S = \sqrt{\frac{1}{k} + \frac{1}{n-k}}$. So

$$H_0 \Rightarrow W = \left(\frac{M - \beta_0}{S}\right) \dot{\sim} N(0,1)$$

$$W = \frac{\hat{\beta} - \beta_0}{S} = \frac{\log\left(\frac{2}{3}\right) - \log\left(\frac{0.5}{0.5}\right)}{\sqrt{\frac{1}{40} + \frac{1}{60}}} = -1.99$$

Again, we reject H_0 at the 5% level.

Exercise 7.4.1
Model and hypotheses as in 7.2.1.

$$H_0 \Rightarrow \frac{U^2}{V} \dot{\sim} \chi^2_1, \qquad \text{where } U = l'(\pi_0), V = -E[l''(\pi_0)]$$

From Example 7.4.1:

$$\frac{U^2}{V} = \frac{(p - \pi_0)^2}{\pi_0(1-\pi_0)/n} = \frac{(0.4 - 0.5)^2}{0.5(1-0.5)/100} = 4$$

Referring to $\chi^2_{1,0.95} = 3.84$, we again reject H_0 at the 5% level.

Practical Exercise

Question 1

Model: $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, 1)$. $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$.
Can write model as $\bar{X} \sim N(\mu, 1/n)$. If we observe $\bar{X} = \bar{x}$, then:

$$l(\mu|\bar{x}) = -\frac{1}{2}\left(\frac{\bar{x}-\mu}{1/\sqrt{n}}\right)^2$$

For LLR test:

$$H_0: \mu = \mu_0 \Rightarrow \; -2llr(\mu_0) = \left(\frac{\bar{x}-\mu_0}{1/\sqrt{n}}\right)^2 \sim \chi_1^2 \Rightarrow \frac{\bar{x}-\mu_0}{1/\sqrt{n}} \sim N(0,1)$$

For Wald test:

$$H_0: \mu = \mu_0 \Rightarrow \frac{M-\mu_0}{S} \sim N(0,1) \Rightarrow \frac{\bar{x}-\mu_0}{1/\sqrt{n}} \sim N(0,1)$$

For score test:

$$H_0: \mu = \mu_0 \Rightarrow \frac{U^2}{V} \sim \chi_1^2, \qquad \text{where } U = l'(\mu_0), V = -E[l''(\mu_0)]$$

$$l'(\mu) = \left(\frac{\bar{x}-\mu}{1/\sqrt{n}}\right)\sqrt{n} = \frac{\bar{x}-\mu}{1/n} \Rightarrow U = l'(\mu_0) = \frac{\bar{x}-\mu_0}{1/n}$$

$$l''(\mu) = -\frac{1}{1/n} = -n \Rightarrow V = -E[l''(\mu_0)] = -E[-n] = n$$

$$\Rightarrow \frac{U^2}{V} = \frac{\left(\frac{\bar{x}-\mu_0}{1/n}\right)^2}{n} = \left(\frac{\bar{x}-\mu_0}{\sqrt{n}/n}\right)^2 = \left(\frac{\bar{x}-\mu_0}{1/\sqrt{n}}\right)^2$$

and since under $H_0 \Rightarrow \frac{U^2}{V} \sim \chi_1^2$, then $\frac{U}{\sqrt{V}} = \left(\frac{\bar{x}-\mu_0}{1/\sqrt{n}}\right) \sim N(0,1)$, which is the same as for the LLR test and the Wald test. This equivalence is because the Normal log-likelihood (ratio) is exactly quadratic.

Question 2

a) $X \sim f(x|\beta), x > 0$ so

$$E[X] = \int_0^\infty x\, f(x)\, dx = \int_0^\infty x\, \beta\, e^{-\beta x}\, dx$$

We can now integrate by parts, using $\int_a^b u \frac{dv}{dx}\, dx = [uv]_a^b - \int_a^b v \frac{du}{dx} dx$ and setting $u = x$; $\frac{dv}{dx} = \beta\, e^{-\beta x}$. Therefore:

$$\frac{du}{dx} = 1; v = -e^{-\beta x}$$

$$\Rightarrow E[X] = \left[-x\, e^{-\beta x}\right]_0^\infty + \int_0^\infty e^{-\beta x} dx = 0 + \left[-\frac{1}{\beta} e^{-\beta x}\right]_0^\infty = -\frac{1}{\beta} e^{-\infty} + \frac{1}{\beta} e^0 = \frac{1}{\beta}$$

b) Suppose there are n patients, and their survival times are random variables $X_1, \dots, X_n$.
Model: $X_1, \dots, X_n \overset{iid}{\sim} f(x|\beta) = \beta \exp(-\beta x)$. $H_0: \beta = \beta_0$ vs $H_1: \beta \neq \beta_0$.

c)

$$L(\beta|\underline{x}) = \prod_{i=1}^{n} f(x_i|\beta) = \prod_{i=1}^{n} \beta e^{-\beta x_i}$$

$$\Rightarrow l(\beta) = \sum_{i=1}^{n} \log\left(\beta e^{-\beta x_i}\right) = \sum_{i=1}^{n} \log \beta - \sum_{i=1}^{n} \beta x_i = n \log \beta - \beta \sum_{i=1}^{n} x_i = n \log \beta - \beta n \bar{x}$$

$$\Rightarrow l'(\beta) = \frac{n}{\beta} - n\bar{x}$$

Get MLE by solving $l'(\hat{\beta}) = 0$ and checking that $l''(\hat{\beta}) < 0$:

$$l'(\hat{\beta}) = \frac{n}{\hat{\beta}} - n\bar{x} = 0 \Rightarrow \hat{\beta} = \frac{1}{\bar{x}}$$

$$l''(\beta) = -\frac{n}{\beta^2} \Rightarrow l''(\hat{\beta}) = -n\bar{x}^2 < 0$$

LLR test statistic:

$$H_0: \beta = \beta_0 \Rightarrow -2\, llr(\beta_0) \dot{\sim} \chi_1^2$$

$$llr(\beta_0) = l(\beta_0) - l(\hat{\beta}) = n \log \beta_0 - \beta_0 n \bar{x} - n \log \hat{\beta} + \hat{\beta} n \bar{x}$$

$$= n \log \beta_0 - \beta_0 n\bar{x} - n \log \frac{1}{\bar{x}} + n = n(\log(\beta_0 \bar{x}) - \beta_0 \bar{x} + 1)$$

$$\Rightarrow \ -2llr(\beta_0) = -2n(\log(\beta_0 \bar{x}) - \beta_0 \bar{x} + 1)$$

d) i) Score test: $H_0 \Rightarrow \frac{U^2}{V} \dot{\sim} \chi_1^2$ where $U = l'(\beta_0), V = -E[l''(\beta_0)]$

$$U = \frac{n}{\beta_0} - n\bar{x}; \ \ V = -E\left[-\frac{n}{\beta_0^2}\right] = \frac{n}{\beta_0^2}$$

$$\Rightarrow \frac{U^2}{V} = \frac{\left(\frac{n}{\beta_0} - n\bar{x}\right)^2}{\frac{n}{\beta_0^2}} = \left(\frac{\beta_0}{\sqrt{n}}\left(\frac{n}{\beta_0} - n\bar{x}\right)\right)^2 = \left(\sqrt{n} - \sqrt{n}\beta_0\bar{x}\right)^2 = n(1 - \beta_0\bar{x})^2$$

ii) Wald test: $H_0 \Rightarrow W^2 = \left(\frac{M - \beta_0}{S}\right)^2 \dot{\sim} \chi_1^2$, where $M = \hat{\beta} = 1/\bar{x}$, and $S^2 = -1/l''(\hat{\beta}) = 1/(n\bar{x}^2)$ (using results from c).

$$\Rightarrow W^2 = \frac{\left(\frac{1}{\bar{x}} - \beta_0\right)^2}{\frac{1}{n\bar{x}^2}} = n\left(\bar{x}\left(\frac{1}{\bar{x}} - \beta_0\right)\right)^2 = n(1 - \beta_0\bar{x})^2$$

Which is the same as the score test statistic.
[Note that we used the square of W to compare more readily with the other two tests.]
Thus in this particular case the score and Wald tests are the same: however, this will not generally be true in non-Normal contexts.

e) Data $x_1, \dots, x_5$ are 0.5, 1, 1.25, 1.5, 0.75, giving $\bar{x} = 1$.
We test the hypothesis $H_0: \beta = 0.5$ vs $H_1: \beta \neq 0.5$.

i) LLR test: $-2llr(\beta_0) = -2 \times 5 \times (\log(0.5 \times 1) - 0.5 \times 1 + 1) = 1.93$. Referring this to $\chi^2_{1,0.95}$, we do not reject H_0 at the 5% level, so there is no evidence at this level that $\beta \neq 0.5$.
We can get the p-value from Stata:

```
. display chi2tail(1, 1.93)
.16475844
```

so p = 0.16 (or you could use Neave tables to show p>0.1).

ii) Score test: $\frac{U^2}{V} = 5 \times (1 - 0.5 \times 1)^2 = 1.25$. Again, referring this to $\chi^2_{1,0.95}$ we do not reject H_0 at the 5% level. From Stata:

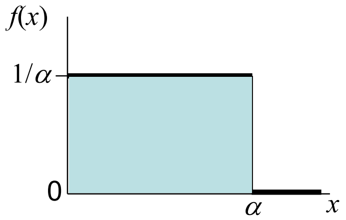
```
. display chi2tail(1, 1.25)
```

```
.26355248
```

we see that the p-value is 0.26.

iii) $W^2 = 1.25$, since it takes the same form as the score test, and again (as we know from the score test) we do not reject. The p-value will also be exactly the same as the score test p-value.

Question 3



$X_1, \dots, X_n \overset{iid}{\sim} \text{Uniform}[0, \alpha]$ (note that this gives the range for the uniform distribution, not the mean and variance!).
$\Rightarrow$ each $X_i \sim f(x|\alpha) = 1/\alpha$, with $0 \leq x \leq \alpha$.
[To see why, note that the shaded area in the figure must have area = 1. Alternatively, $F(x) = \text{Prob}(X < x) = \frac{x}{\alpha} \Rightarrow F'(x) = f(x) = \frac{1}{\alpha}$]
So

$$L(\alpha|\underline{x}) = \prod_{i=1}^{n} f(x_i|\alpha) = \prod_{i=1}^{n} \frac{1}{\alpha} = \frac{1}{\alpha^n} \text{ if } 0 \leq x_i \leq \alpha$$

with $L(\alpha) = 0$ if *any* $x_i > \alpha$ or $x_i < 0$.

Thus the whole likelihood is set to zero if any x_i is greater than α (since the true value of the parameter has to 'contain' all the data), otherwise it equals $1/\alpha^n$. So the maximum of the likelihood is at the smallest value of α such that all $x_i \leq \alpha$, therefore $\hat{\alpha} = \max(x_i)$. The likelihood function looks as in the figure below, with the MLE $\hat{\alpha} = \max(x_i)$ at a discontinuity in the likelihood: consequently the maximum likelihood estimate properties will not apply and we will have problems testing hypotheses.