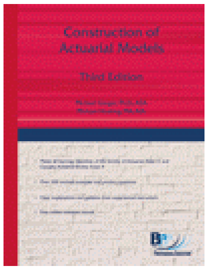

Construction of Actuarial Models

Third Edition

by Mike Gauger and Michael Hosking

Published by BPP Professional Education

Solutions to practice questions – Chapter 6

Solution 6.1

With 6 degrees of freedom, the 95th percentile is 12.592 and the 97.5th percentile is 14.449. Since the statistic value 13.5 is between these two percentiles we can say that the p-value is between 2.5% and 5%.

Solution 6.2

The null and alternative hypotheses are:

H_0 : the Binomial $(m=4, q=0.5)$ model is correct

H_1 : the Binomial $(m=4, q=0.5)$ model is not correct

Under the suggested model the expected frequencies are:

Number of boys	0	1	2	3	4	Total
Expected frequency	9.375	37.5	56.25	37.5	9.375	150

These figures have been calculated using the formula:

$$150\Pr(N=n) = 150\binom{4}{n}0.5^n\,0.5^{4-n} = 150\binom{4}{n}0.5^4$$

The test statistic is:

$$\chi^2 = \frac{(8-9.375)^2}{9.375} + \frac{(36-37.5)^2}{37.5} + \frac{(61-56.25)^2}{56.25} + \frac{(32-37.5)^2}{37.5} + \frac{(13-9.375)^2}{9.375}$$

$$= 2.871$$

We compare this with χ_4^2. The upper 5% point of χ_4^2 is 9.49. As the test statistic is less than this critical value, we have insufficient evidence to reject the null hypothesis and we conclude that the binomial model is correct.

Solution 6.3

We would have compared the value of the test statistic with χ_3^2, instead of χ_4^2.

Solution 6.4

The following values are realizations of a random variable Y:

100 150 225 290 300 500

You want to test whether these data come from a Pareto distribution with parameters $\alpha = 3$ and $\theta = 600$.

The hypothesized CDF is:

$$F(y) = 1 - \left(\frac{600}{y+600}\right)^3$$

For the given data we have:

j	$y_{(j)}$	$F\left(y_{(j)}\right)$	$\frac{j}{6} - F\left(y_{(j)}\right)$	$F\left(y_{(j)}\right) - \frac{j-1}{6}$
1	100	0.370	–0.204	0.370
2	150	0.488	–0.155	0.321
3	225	0.615	–0.115	0.282
4	290	0.694	–0.027	0.194
5	300	0.703	0.130	0.037
6	500	0.838	0.162	0.004

So the Kolmogorov-Smirnov test statistic is 0.370.

Solution 6.5

The geometric distribution is a special case of the negative binomial distribution with the restriction that $r = 1$. As we are given the maximized log-likelihood values, the likelihood ratio test statistic is:

$$-2(-578.9 + 573.6) = 10.6$$

Solution 6.6

Using the suggested percentiles, we have the following results:

Range	Observed	Expected
0 - 1,000	3	4
1,000 - 3,000	7	8
3,000 - 5,000	10	6
5,000 - 8,000	9	10
8,000 - 11, 000	5	8
11,000 +	6	4

Combining the first two groups and the last two groups to make the expected values at least 5, we obtain the following test statistic:

$$\chi^2 = \frac{(10-12)^2}{12} + \frac{(10-6)^2}{6} + \frac{(9-10)^2}{10} + \frac{(11-12)^2}{12} = 3.183$$

 © BPP Professional Education

Solution 6.7

We first have to calculate the cdf. If $0 < x \le ½$, then:

$$F(x) = \int_0^x 4t\,dt = \left[2t^2\right]_0^x = 2x^2$$

If $½ < x < 1$, then:

$$F(x) = ½ + \int_{½}^x (4-4t)\,dt = ½ + \left[4t - 2t^2\right]_{½}^x = ½ + \left(4x - 2x^2\right) - (2 - ½) = 4x - 2x^2 - 1$$

So we have:

j	$x_{(j)}$	$F(x_{(j)})$	$\frac{j}{4} - F(x_{(j)})$	$F(x_{(j)}) - \frac{j-1}{4}$
1	0.16	0.0512	0.1988	0.0512
2	0.39	0.3042	0.1958	0.0542
3	0.62	0.7112	0.0388	0.2112
4	0.75	0.8750	0.1250	0.1250

The maximum of the numbers in the last two columns is 0.2112, and this is the Kolmogorov-Smirnov test statistic.

Solution 6.8

The ranking criterion is $\ln L - \frac{r}{2}\ln n$, where r is the number of parameters in the model and n is the number of data points. The figures are as follows:

Inverse Burr model: $\ln(L) - \frac{r}{2}\ln(n) = -180.2 - \frac{3}{2}\ln(200) = -188.1$

Inverse Gamma model: $\ln(L) - \frac{r}{2}\ln(n) = -181.4 - \frac{2}{2}\ln(200) = -186.7$

Weibull model: $\ln(L) - \frac{r}{2}\ln(n) = -181.6 - \frac{2}{2}\ln(200) = -186.9$

Exponential model: $\ln(L) - \frac{r}{2}\ln(n) = -183.0 - \frac{1}{2}\ln(200) = -185.6$

The largest of these numbers is –185.6, so the best model according to the Schwarz-Bayesian criterion is the Exponential.

Solution 6.9

The observation 267 is the sixth out of seven when they are put in increasing order. The 2-parameter Pareto CDF is:

$$F(x) = 1 - \left(\frac{\theta}{\theta + x}\right)^{\alpha} = 1 - \left(\frac{250}{250 + x}\right)^2$$

So the point sought is:

$$\left(\frac{6}{7+1}\ ,\ F(267)\right) = (0.750\ ,\ 0.766)$$

Solution 6.10

(i) $f(x) = \dfrac{\beta^x}{(1+\beta)^{x+1}}$ for $x = 0, 1, \ldots \Rightarrow L(\beta) = \dfrac{\beta^{\sum x_i}}{(1+\beta)^{\sum (x_i+1)}}$

$$\Rightarrow \ln\big(L(\beta)\big) = \sum x_i \times \ln(\beta) - \sum (x_i+1) \times \ln(1+\beta)$$

$$\Rightarrow 0 = \frac{d\ln\big(L(\beta)\big)}{d\beta} = \frac{\sum x_i}{\beta} - \frac{\sum (1+x_i)}{1+\beta} = \frac{9}{\beta} - \frac{9+8}{1+\beta} \Rightarrow \hat{\beta} = \frac{9}{8}$$

(ii) The maximum value is:

$$\ln\left(L\left(\hat{\beta}\right)\right) = \sum x_i \times \ln(9/8) - \sum (x_i+1) \times \ln(1+9/8)$$
$$= 9 \times \ln(9/8) - 17 \times \ln(17/8) = -11.754$$

(iii) There is 1 restriction. The 95th percentile of the chi-squared distribution with 1 degree of freedom is 3.841.

So the null hypothesis is rejected if the statistic exceeds 3.841:

$$3.841 < 2\big(\ln(L_{UR}) - \ln(L_R)\big) = 2\big(\ln(L_{UR}) + 11.754\big) \Leftrightarrow -9.834 < \ln(L_{UR})$$

 © BPP Professional Education