ACSC/STAT 4703, Actuarial Models II

FALL 2022

Toby Kenney

Midterm Examination

Model Solutions

Here are some values of the Gamma distribution function with $\theta = 1$ that may be needed for this examination:

x	α	$F(x)$	x	α	$F(x)$	x	α	$F(x)$
245	255	0.2697208	2.5	4	0.2424239	4.375	4	0.6361773
$\left(\frac{7.5}{12}\right)^3$	$\frac{4}{3}$	0.1117140	3.841	2.4	0.8409823	4.875	4	0.7169870
$\left(\frac{9.5}{12}\right)^3$	$\frac{4}{3}$	0.2507382	4.375	3	0.8118663	5.375	4	0.7837292
1.356	2.4	0.2801616	4.875	3	0.8644174	2.156	5	0.06782354
1.941	2.4	0.4612472	5.375	3	0.9035828	3.203	5	0.219922
2.367	2.4	0.5775816	3.875	4	0.5417358	8.542	5	0.9274742

Here are the critical values for a chi-squared distribution:

Degrees of Freedom	Significance level 90%	95%	99%
1	2.705543	3.841459	6.634897
2	4.605170	5.991465	9.210340
3	6.251389	7.814728	11.344867
4	7.779440	9.487729	13.276704
5	9.236357	11.070498	15.086272

1. *Using an arithmetic distribution ($h = 1$) to approximate an inverse Weibull distribution with $\tau = 2$ and $\theta = 6$, calculate the probability that the value is more than 8.5, for the approximation using the method of local moment matching, matching 1 moment on each interval.*

   [Hint:

$$\int_8^9 x^{-2}e^{-\left(\frac{6}{x}\right)^2}\,dx = 0.00840944$$

   ]

   If we let r_n and s_n be the probability assigned to n from the intervals $[n-1, n]$ and $[n, n+1]$ respectively, so that $p_n = r_n + s_n$, then we have $p_1 + \ldots + p_7 + r_8 = F_X(8) = e^{-\left(\frac{6}{8}\right)^2} = 0.569782824731$. We have

$$\begin{aligned}
s_8 + r_9 &= e^{-\left(\frac{6}{9}\right)^2} - e^{-\left(\frac{6}{8}\right)^2} = 0.071397563699\\
8s_8 + 9r_9 &= \int_8^9 2\frac{6^2}{x^2}e^{-\left(\frac{6}{x}\right)^2}\,dx = 72 \times 0.00840944 = 0.60547968\\
s_8 &= 9 \times 0.071397563699 - 0.60547968\\
&= 0.037098393291
\end{aligned}$$

Thus $P(X_a > 8.5) = 1 - 0.569782824731 - 0.037098393291 = 0.393118781978$.

2. *Claim frequency follows a negative binomial distribution with* $r = 3.3$ *and* $\beta = 1.5$. *Claim severity (in thousands) has the following distribution:*

Severity	*Probability*
0	*0.46*
1	*0.33*
2	*0.16*
$\geqslant 3$	*0.05*

The expected claim severity per loss is 0.81. *The company buys excess-of loss reinsurance for aggregate losses exceeding 2.*

(a) Use the recursive method to calculate the probability that the reininsurance makes a payment.

For the negative binomial distribution, we have $a = \frac{\beta}{1+\beta} = 0.6$ and $b = (r-1)\frac{\beta}{1+\beta} = 1.38$. The recurrence relation is therefore

$$f_S(x) = \frac{\sum_{y=1}^{x} \left(0.6 + 1.38\frac{y}{x}\right) f_X(y) f_S(x-y)}{1 - 0.6 f_X(0)}$$

We also have

$$f_S(0) = P_N(f_X(0)) = (2.5 - 1.5 \times 0.46)^{-3.3} = 0.141143441599$$

We calculate

$$f_S(1) = \frac{1.98 \times 0.33 \times 0.141143441599}{0.724} = 0.127380006548$$
$$f_S(2) = \frac{1.29 \times 0.33 \times 0.127380006548 + 1.98 \times 0.16 \times 0.141143441599}{0.724} = 0.136657335754$$

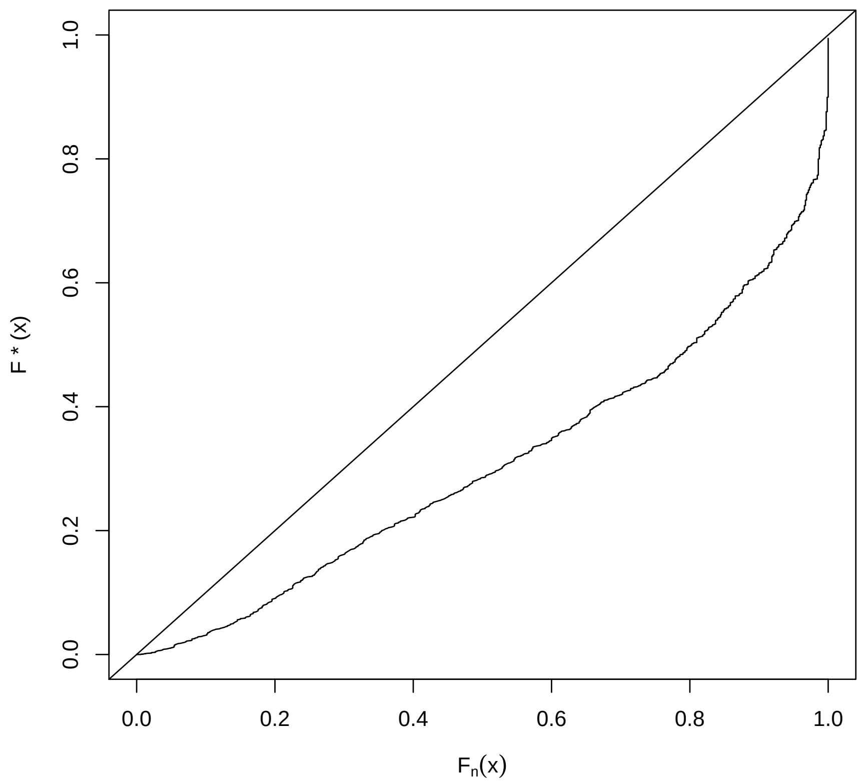
Thus the probability that the reinsurance company makes a payment is $1 - 0.141143441599 - 0.127380006548 - 0.136657335754 = 0.594819216099$.

(b) What is the expected payment on the reinsurance? [Hint: first calculate the insurer's expected payment with this reinsurance policy. Then consider the expected total payments between the insurer and the reinsurer.]

The insurer pays 0 if $S = 0$, 1 if $S = 1$ and 2 if $S \geqslant 2$. Thus the insurer's expected payment is $f_S(1) + 2(1 - f_S(0) - f_S(1)) = 2 - 0.127380006548 - 2 \times 0.141143441599 = 1.59033311025$.

The expected number of losses is $3.3 \times 1.5 = 4.95$, and the expected claim per loss is 0.81, so the expected aggregate claim is $4.95 \times 0.81 = 4.0095$. The insurer pays 1.59033311025 of this, so the reinsurer's expected payment is $4.0095 - 1.59033311025 = 2.41916688975$.

3. *An insurance company collects a sample of 700 claims. Based on previous experience, it believes these claims might follow an inverse Pareto distribution with* $\theta = 0.7$ *and* $\tau = 3.9$. *To test this, it computes the following p-p plot.*

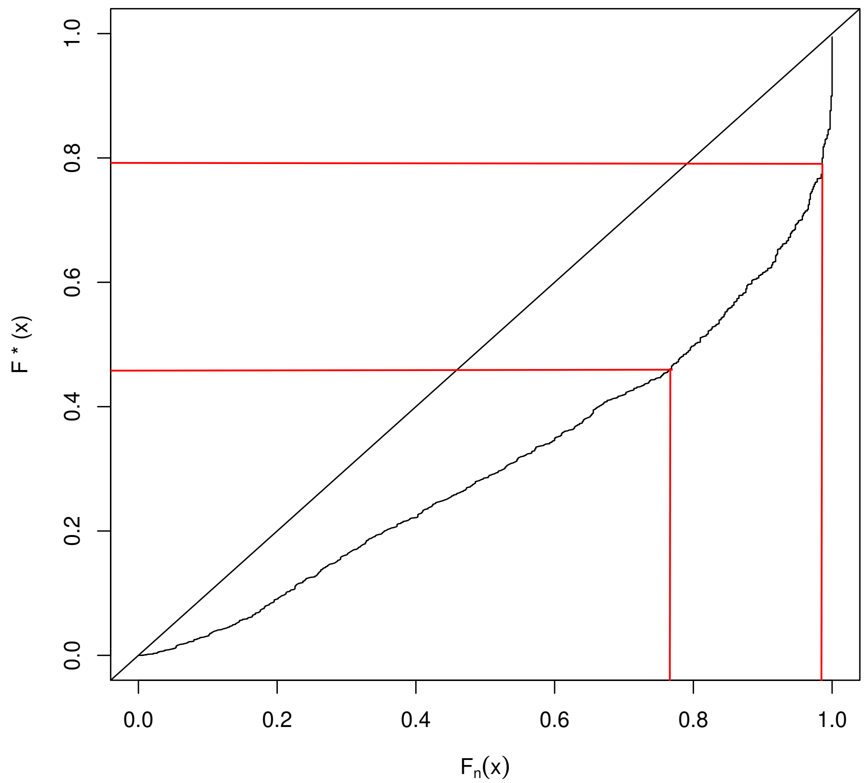


(a) How many of the claims in their sample were between 3 and 11?

We have that

$$F^*(3) = \left(\frac{3}{3+0.7}\right)^{3.9} = 0.441353087998$$
$$F^*(11) = \left(\frac{11}{11+0.7}\right)^{3.9} = 0.786152158728$$

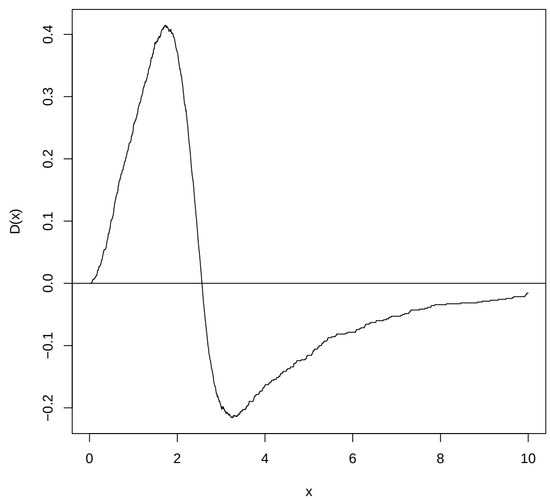
From the graph:



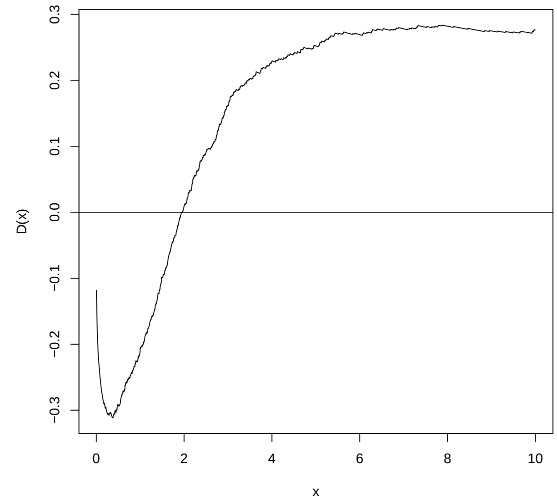
we read that $F_n(3) \approx 0.76$ and $F_n(11) \approx 0.98$. Since there are 700 samples, there are approximately $(0.98 - 0.76) \times 700 = 154$ samples between 3 and 11 in the dataset.

[There are actually 174 samples between 3 and 11 in the dataset.]

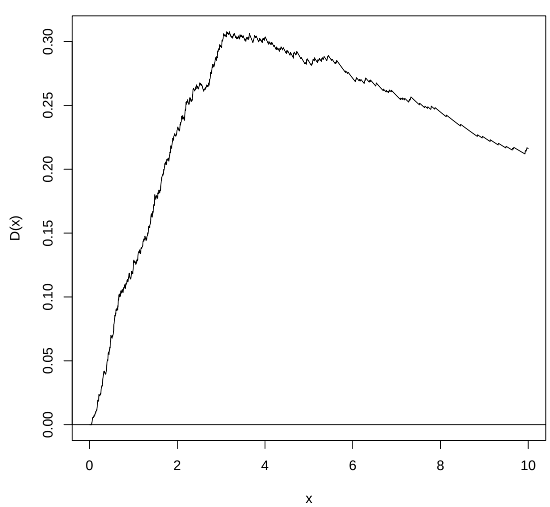
(b) Which of the following is a plot of $D(x) = F_n(x) - F^*(x)$ *for this data?*



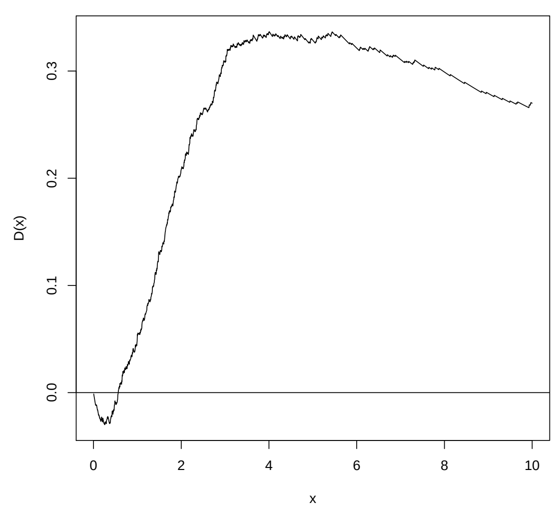
(i)



(ii)



(iii)



(iv)

Justify your answer.

From the p-p plot, we see that $F^*(x) < F_n(x)$ for all x, so $D(x)$ should be positive for all x. This is only the case for (iii). Therefore (iii) is the true plot of $D(x)$.

4. *An insurance company collects the following sample:*

0.06 0.32 0.61 0.67 1.16 2.53 4.02 5.09 10.27 15.83 17.64 17.84 20.00 20.92 24.44 42.52 63.80 71.84

They model this as following a distribution with the following distribution function:

x	$F(x)$	$i^2(\log(F(x_{i+1})) - \log(F(x_i)))$	$(18-i)^2(\log(1-F(x_i)) - \log(1-F(x_{i+1})))$
0.06	0.003940665	3.6581388	52.466840
0.32	0.152854776	1.3462645	21.656703
0.61	0.214016777	0.8161002	6.703294
0.67	0.234330430	5.0055286	26.831221
1.16	0.320402414	1.5140605	5.855879
2.53	0.340406390	4.2591188	11.325024
4.02	0.383158609	13.8026392	32.511770
5.09	0.507824310	1.0051731	1.992625
10.27	0.515863074	2.8523340	3.893844
15.83	0.534352306	9.9014547	10.302913
17.64	0.589968955	12.5596075	10.958980
17.84	0.654497867	17.0641094	13.339362
20.00	0.736838758	13.0086484	9.131373
20.92	0.795796371	5.5330811	2.957822
24.44	0.818581810	22.8721986	10.549221
42.52	0.906170573	4.0967235	1.524215
63.80	0.920788516	6.6446792	1.260842
71.84	0.942204512	19.2886673	
		145.228528	223.261928

Calculate the Anderson-Darling statistic for this model and this data.

The Anderson-Darling statistic for this data is

$$-nF^*(u) + n\sum_{i=1}^{k}(F_n(y_i))^2\left(\log(F^*(y_{i+1})) - \log(F^*(y_i)\right) + n\sum_{i=0}^{k}(1-F_n(y_i))^2\left(\log(1-F^*(y_i)) - \log(1-F^*(y_{i+1})\right)$$
$$= n\left(\sum_{i=1}^{k}\frac{i^2}{n^2}\left(\log(F^*(y_{i+1})) - \log(F^*(y_i)\right) + \sum_{i=0}^{k}\frac{(n-i)^2}{n^2}\left(\log(1-F^*(y_i)) - \log(1-F^*(y_{i+1})\right) - 1\right)$$
$$= 18\left(\frac{145.228528}{18^2} + \frac{223.261928}{18^2} - 1\right)$$
$$= 2.47169199996$$

5. *An insurance company collects a sample of 1500 claims. They want to decide whether this data is better modeled as following an inverse exponential distribution, or a generalised Pareto distribution. After calculating MLE estimates for the parameters (1 parameter for the inverse exponential and 3 for the generalised Pareto), log-likelihoods for the two distributions are:*

Distribution	log-likelihood
Inverse Exponential	-4244.75
Generalised Pareto	-4236.89

Use a BIC to decide whether the generalised Pareto distribution or the inverse exponential distribution is a better fit for the data.

The BIC for the inverse exponential is $-4244.75 - \frac{1}{2}\log(1500) = -4248.40661019$. The BIC for the Generalised Pareto is $-4236.89 - \frac{3}{2}\log(1500) = -4247.85983058$. Therefore, the Generalised Pareto is prefered.

6. *A homeowner's house has a value of $860,000. The insurer requires 75% coverage for full insurance. The deductible is $4,000, decreasing linearly to zero for losses of $12,000. The home sustains $7,000 of damage from fire. The insurer reimburses $3,300. For what value was the home insured?*

   The deductible for a loss of $7,000 is $\frac{12000-7000}{12000-4000} \times 4000 = \$2,500.$. Thus, under, full insurance, the insurer would reimburse $7000 - 2500 = \$4,500$. Therefore, the coverage for this home is $\frac{3300}{4500} = \frac{11}{15}$. For full insurance, the home should be insured for $860000 \times 0.75 = \$645,000$. Therefore, the home is insured for $\frac{11}{15} \times 645000 = \$473,000$.

7. *The following table shows the cumulative losses (in thousands) on claims from one line of business of an insurance company over the past 4 years.*

Accident year	*Development year* 0	*1*	*2*	*3*
2018	*5539*	*6003*	*6829*	*7108*
2019	*6243*	*6792*	*7314*	
2020	*6217*	*7209*		
2021	*6372*			

*Using the mean for calculating loss development factors, esimate the total reserve needed for payments to be made in **2023** (two years in the future) using the Bornhuetter-Fergusson method. The expected loss ratio is 0.74 and the earned premiums in each year are given in the following table:*

Year	*Earned Premiums (000's)*
2018	*8537*
2019	*8764*
2020	*9023*
2021	*9285*

[Assume no more payments are made after development year 3.]

The mean loss development factors are

Development Year	Loss Development Factor
0/1	$\frac{20004}{17999} = 1.1113950775$
1/2	$\frac{14143}{12795} = 1.10535365377$
2/3	$\frac{7108}{6829} = 1.04085517645$

The proportion of cumulative payments by the end of each development year are:

Development Year	Cumulative Proportion of payments made	Proportion of payments made
0	$\frac{1}{1.1113950775\times1.10535365377\times1.04085517645} = 0.782059819772$	0.782059819772
1	$\frac{1}{1.10535365377\times1.04085517645} = 0.869177434008$	0.087117614236
2	$\frac{1}{1.04085517645} = 0.960748452451$	0.091571018443
3	1	0.039251547549

The expected payments to be made in 2023 from Accident year 2020 are $9023\times0.74\times0.039251547549 = 262.083368016$.
The expected payments to be made in 2023 from Accident year 2021 are $9285\times0.74\times0.091571018443 = 629.17531062$.
Thus, the total reserves needed for payments to be made in 2023 are $262.083368016 + 629.17531062 = \$891.26$.