ACSC/STAT 4703, Actuarial Models II

FALL 2023
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	Significance level		
Freedom	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 a Generalised Pareto distribution with $\xi = -4$ and $\beta = 50$, calculate the probability that the value is more than 4.5, for the approximation using the method of local moment matching, matching 1 moment on each interval.*

   Under the Generalised Pareto distribution, the probability of the interval $[n, n+1]$ is $\left(1 - 4\frac{n}{50}\right)^{\frac{1}{4}} - \left(1 - 4\frac{n+1}{50}\right)^{\frac{1}{4}}$, and the probability multiplied by the conditional mean is

$$\int_n^{n+1} 0.02x(1 - 0.08x)^{-\frac{3}{4}}\,dx = \int_{0.92-0.08n}^{1-0.08n} 0.25 \times 12.5(1-u)u^{-\frac{3}{4}}\,du = 3.125\int_{0.92-0.08n}^{1-0.08n} u^{-\frac{3}{4}} - u^{\frac{1}{4}}\,du = 3.125\left[4u^{\frac{1}{4}} - \frac{4}{5}u^{\frac{5}{4}}\right]_{0.92-0.08n}^{1-0.08n}$$

   In particular, for $n = 4$, this is

$$3.125\left[4u^{\frac{1}{4}} - \frac{4}{5}u^{\frac{5}{4}}\right]_{0.6}^{0.68} = 12.5\left(0.68^{\frac{1}{4}} - 0.6^{\frac{1}{4}}\right) - 2.5\left(0.68^{\frac{5}{4}} - 0.6^{\frac{5}{4}}\right) = 0.12610529532$$

   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 $p_1 + \ldots + p_3 + r_4 = F_X(4) = 1 - (1 - 0.08 \times 4)^{\frac{1}{4}} = 0.091913481477$. We have

$$\begin{aligned} s_4 + r_5 &= (1 - 0.08 \times 4)^{\frac{1}{4}} - (1 - 0.08 \times 5)^{\frac{1}{4}} = 0.027974781729 \\ 4s_4 + 5r_5 &= 0.12610529532 \\ s_4 &= 5 \times 0.027974781729 - 0.12610529532 \\ &= 0.013768613325 \end{aligned}$$

Thus $P(X_a > 4.5) = 1 - 0.091913481477 - 0.013768613325 = 0.894317905198$.

15 mins Cumulative Time: 15

2. *Claim frequency follows a Poisson distribution with* $\lambda = 3.5$. *Claim severity (in thousands) has the following distribution:*

Severity	*Probability*
0	*0.62*
1	*0.24*
2	*0.07*
$\geqslant 3$	*0.07*

The expected claim severity per loss is 0.58. *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 Poisson distribution, we have $a = 0$ and $b = \lambda = 3.5$. The recurrence relation is therefore

$$f_S(x) = \sum_{y=1}^{x} 3.5\frac{y}{x} f_X(y) f_S(x-y)$$

We also have

$$f_S(0) = P_N(f_X(0)) = e^{3.5(0.62-1)} = 0.2644772613$$

We calculate

$$\begin{aligned} f_S(1) &= 3.5 \times 0.24 \times 0.2644772613 = 0.222160899492 \\ f_S(2) &= 3.5 \times \frac{1}{2} \times 0.24 \times 0.222160899492 + 3.5 \times 0.07 \times 0.2644772613 = 0.158104506805 \end{aligned}$$

Thus the probability that the reinsurance company makes a payment is $1 - 0.2644772613 - 0.222160899492 - 0.158104506805 = 0.355257332403$.

10 mins Cumulative Time: 25

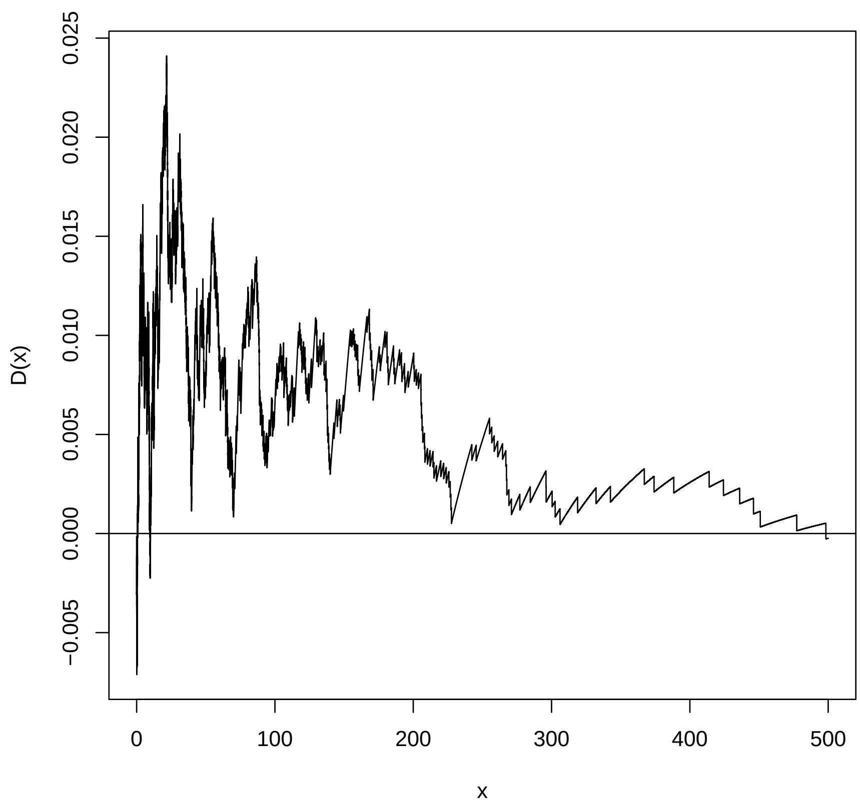
(b) What is the expected payment on the reinsurance? [Hint: first calculate the insurer's expected payment with this reinsurance policy. Then consider the expected total payments between the insurer and the reinsurer.]

The insurer pays 0 if $S = 0$, 1 if $S = 1$ and 2 if $S \geqslant 2$. Thus the insurer's expected payment is $f_S(1) + 2(1 - f_S(0) - f_S(1)) = 2 - 0.222160899492 - 2 \times 0.2644772613 = 1.24888457791$.

The expected number of losses is 3.5, and the expected claim per loss is 0.81, so the expected aggregate claim is $3.5 \times 0.58 = 2.03$. The insurer pays 1.24888457791 of this, so the reinsurer's expected payment is $2.03 - 1.24888457791 = 0.78111542209$.

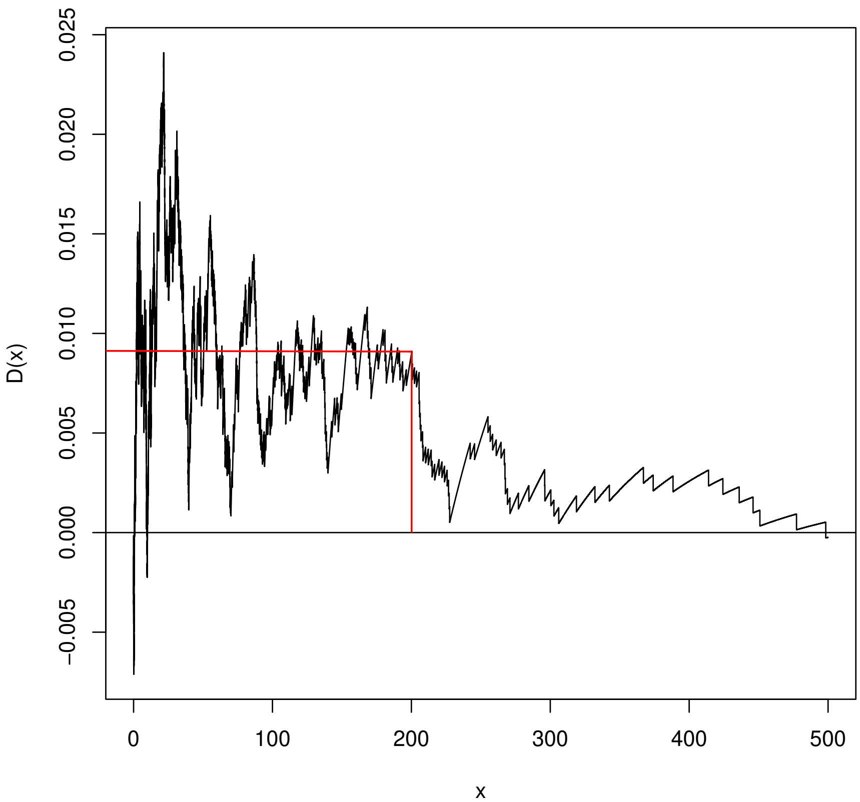
10 mins Cumulative Time: 35

3. *An insurance company collects a sample of 1265 claims. Based on previous experience, it believes these claims might follow a Weibull distribution with $\theta = 36$ and $\tau = 0.7$. To test this, it computes the following plot of $D(x) = F^*(x) - F_n(x)$.*



(a) How many of the claims in their sample were more than 200?

From the graph, we read $D(200) \approx 0.009$

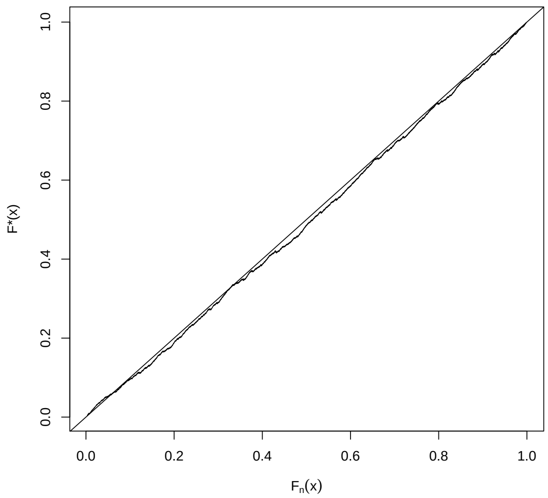


We have $F^*(200) = 1 - e^{-\left(\frac{200}{36}\right)^{0.7}} = 0.96389445143$, so $F_n(200) = 0.96389445143 - 0.009 = 0.95489445143$. Since there are 1265 samples, there are approximately $(1 - 0.95489445143) \times 1265 = 57$ samples more than 200 in the dataset.

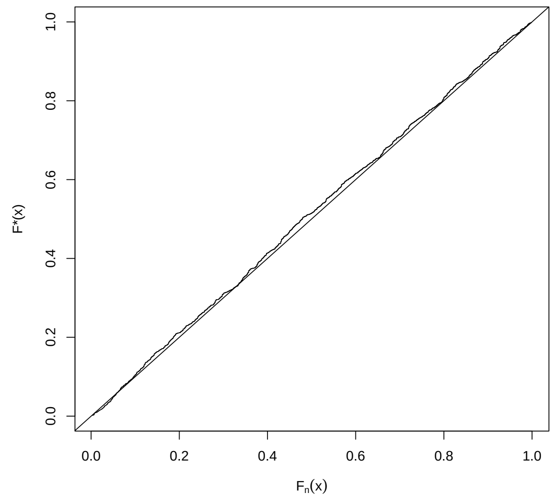
[There are actually 57 samples more than 200 in the dataset.]

5 mins Cumulative Time: 40

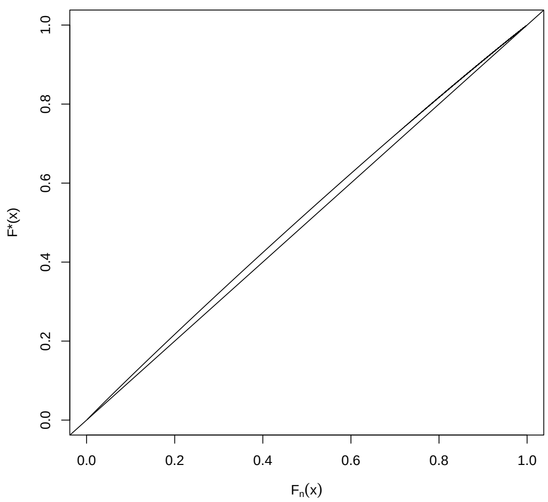
(b) Which of the following is a p-p plot of this data?



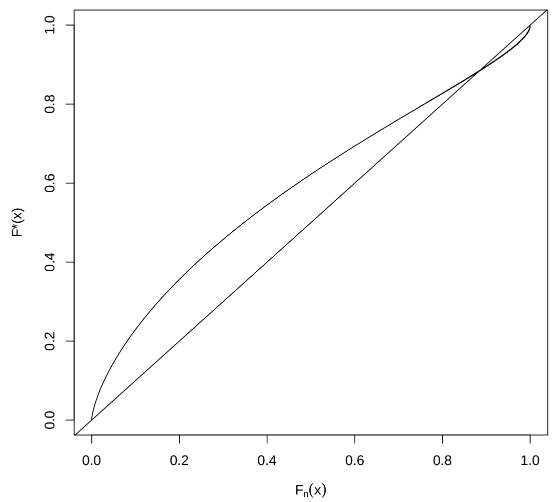
(i)



(ii)



(iii)



(iv)

Justify your answer.

From the plot of $D(x)$, we see that $F^*(x) - F_n(x) > 0$ for nearly all x, so the *p-p* plot should lie above the line $y = x$ for almost the whole graph. This rules out (i). For the others from the plot of $D(x)$, we see the largest value of $D(x)$ is about 0.025, and occurs at about $x = 25$, where $F^*(x) = 1 - e^{-\left(\frac{25}{36}\right)^{0.7}} = 0.53916842692$. We see that in (iv), the largest value of $D(x)$ is much larger than this. In (iii), the largest value of $D(x)$ is approximately 0.25, but $D(x)$ is a lot smoother than in the original plot. Also, we note that $D(x)$ is negative for very small x, which we do not see in (iii), so (ii) is the correct plot.

10 mins Cumulative Time: 50

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

`0.13 0.23 0.23 0.27 0.59 1.26 1.28 1.76 2.33 5.04 6.89 8.16 9.09 13.86 15.92 16.89`

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

i	x_i	$F(x)$	$i^2(\log(F(x_{i+1})) - \log(F(x_i)))$	$(16-i)^2(\log(1-F(x_i)) - \log(1-F(x_{i+1})))$
0		0		37.43461
1	0.13	0.1360401	0.106478454	4.56953483
2	0.23	0.1513248	0.002172436	0.02179601
3	0.23	0.1514070	4.050981349	20.96220820
4	0.27	0.2374793	6.400518565	28.10090523
5	0.59	0.3542889	0.669214036	2.15965902
6	1.26	0.3639008	2.301255097	4.65792993
7	1.28	0.3879223	0.980911115	1.28980418
8	1.76	0.3957662	2.119481049	1.80639628
9	2.33	0.4090922	12.271987807	7.69244240
10	5.04	0.4760137	7.227572221	3.45541285
11	6.89	0.5116917	16.664763579	6.05162201
12	8.16	0.5872482	15.074621810	4.27021623
13	9.09	0.6520573	16.286311186	3.36337902
14	13.86	0.7180227	4.404637982	0.53652559
15	15.92	0.7343413	60.369489479	7.60702336
16	16.89	0.9603355	10.360980876	
Total			159.291377	111.020333

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

The Anderson-Darling statistic for this data is

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

$$= n\left(\sum_{i=1}^{k}\frac{i^2}{n^2}\left(\log(F^*(y_{i+1})) - \log(F^*(y_i)) + \sum_{i=0}^{k}\frac{(n-i)^2}{n^2}\left(\log(1-F^*(y_i)) - \log(1-F^*(y_{i+1})) - 1\right)\right.\right.$$

$$= 16\left(\frac{159.291377}{16^2} + \frac{111.020333}{16^2} - 1\right)$$

$$= 16\left(\frac{159.291377}{16^2} + \frac{111.020333}{16^2} - 1\right)$$

$$= 0.89448187504$$

10 mins Cumulative Time: 60

5. *An insurance company collects a sample of 1952 claims. They want to decide whether this data is better modelled as following a Gumbel distribution or a Fréchet distribution. After calculating MLE estimates for the parameters (2 parameters for the Gumbel and 3 for the Fréchet), log-likelihoods for the two distributions are:*

Distribution	*log-likelihood*
Gumbel	*-5049.35*
Fréchet	*-5048.89*

Use AIC to decide whether the Gumbel distribution or the Fréchet distribution is a better fit for the data.

The AIC for the Gumbel is $-5049.35 - 1 = -5050.35$. The AIC for the Fréchet is $-5048.89 - 2 = -5050.89$. Therefore, the Gumbel distribution is prefered.

5 mins Cumulative Time: 65

6. *An insurer's premium for a policy with limit $1,000,000 is the expected loss plus a risk charge equal to the square of the expected loss divided by $4,000. The pure premium ILF from $1,000,000 to $2,000,000 is 1.17. By buying excess-of-loss reinsurance with attachment point $1,000,000 and limit $1,000,000, and loading 25%, the insurer is able to charge a premium of $880 for policies with limit $2,000,000. What is the premium for a policy with limit $1,000,000?*

Let x be the expected loss with limit $1,000,000 and y be the expected loss with limit $2,000,000. The pure premium ILF of 1.17 means that $y = 1.17x$ The premium with limit $1,000,000 is $x + \frac{x^2}{4000}$. After buying the reinsurance with 25% loading, the premium with limit $2,000,000 is $x + \frac{x^2}{4000} + 1.25(y - x) = 1.2125x + \frac{x^2}{4000}$, so we are given that

$$1.2125x + \frac{x^2}{4000} = 880$$

which we can solve to get

$$x = \frac{\sqrt{4850^2 + 14080000} - 4850}{2} = 641.04386792$$

The premium for the policy with limit $1,000,000 is therefore

$$641.04386792 + \frac{641.04386792^2}{4000} = \$743.78$$

10 mins Cumulative Time: 75