ACSC/STAT 4703, Actuarial Models II
Fall 2017
Toby Kenney
Homework Sheet 2
Due: Friday 6th October: 11:30 PM

Basic Questions

1. An insurance company has the following portfolio of workers compensation insurance policies:

Type of worker	Number	Probability of claim	mean claim	standard deviation
Manual labourer	1700	0.01	\$54,000	\$129,000
Technician	800	0.002	\$20,000	\$39,000
Manager	200	0.001	\$25,000	\$20,000

   Calculate the cost of reinsuring losses above \$10,000,000, if the loading on the reinsurance premium is one standard deviation above the expected claim payment on the reinsurance policy using a gamma approximation for the aggregate losses on this portfolio.

2. An insurance company is modelling claim data as following a Pareto distribution with $\alpha = 5$. It collects the following sample of claims:

```
24.2 26.5 56.9 68.9 72.3 116.8 128.5 145.6 151.3 173.9
181.8 189.4 206.4 229.3 243.3 273.6 303.7 344.0 367.0
375.0 378.5 465.4 500.9 633.9 635.1 638.6 641.9 748.7
2047.2 2895.9
```

   The MLE for θ is 1744.23679. Graphically compare this empirical distribution with the best fitting Pareto distribution with $\alpha = 5$. Include the following plots:

   (a) Comparisons of $F(x)$ and $F^*(x)$

   (b) Comparisons of $f(x)$ and $f^*(x)$

   (c) A plot of $D(x)$ against x.

   (d) A p-p plot of $F(x)$ against $F^*(x)$.

3. For the data in Question 2, calculate the following test statistics for the goodness of fit of the Pareto distribution with $\alpha = 5$ and $\theta = 1744.23679$:

   (a) The Kolmogorov-Smirnov test.

   (b) The Anderson-Darling test.

   (c) The chi-square test, dividing into the intervals 0–200, 200–400, and more than 400.

4. For the data in Question 2, perform a likelihood ratio test to determine whether a Pareto distribution with fixed $\alpha = 5$, or a Pareto distribution with α freely estimated is a better fit for the data. [The MLE for the general Pareto distribution is $\alpha = 4.641528$ and $\theta = 1599.8973$.]

5. For the data in Question 2, use AIC and BIC to choose between a Pareto distribution with $\alpha = 4$ and a gamma distribution for the data. [The MLE for the gamma distribution is $\alpha = 1.021439$ and $\theta = 432.8697$.]

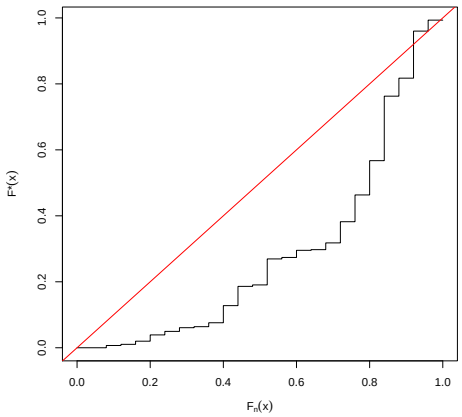
Standard Questions

6. An insurance company insures drivers in three provinces and has the following estimates:

Province	Probability of claim	mean claim	standard deviation
Nova Scotia	0.08	\$3,200	\$5,900
New Brunswick	0.03	\$2,100	\$3,400
PEI	0.02	\$2,300	\$4,600

   The insurance company estimates the mean μ and standard deviation σ for the aggregate loss distribution, and buys stop-loss reinsurance for aggregate losses more than $3\mu - \frac{\mu^2}{\sigma}$. The reinsurer models aggregate losses as following a Pareto distribution and sets its premium as 125% of the expected claims on the stop-loss policy. The insurer already insures 2,200 drivers in Nova Scotia, and 980 drivers in New Brunswick. How many drivers should it insure in PEI in order to minimise the reinsurance cost as a proportion of expected claims on the policy?

7. An insurance company collects a sample of 25 past claims, and attempts to fit a distribution to the claims. Based on experience with other claims, the company believes that an exponential distribution with mean $\theta = 2,400$ may be appropriate to model these claims. It constructs the following p-p plot to compare the sample to this distribution:



(a) How many of the points in their sample were less than 3,000?

(b) Which of the following statements best describes the fit of the exponential distribution to the data:

(i) The exponential distribution assigns too much probability to high values and too little probability to low values.

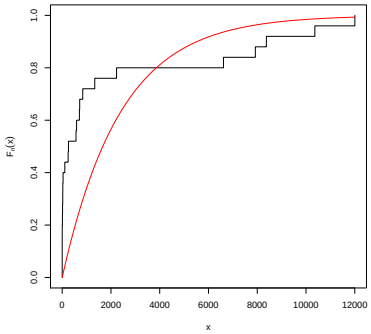
(ii) The exponential distribution assigns too much probability to low values and too little probability to high values.

(iii) The exponential distribution assigns too much probability to tail values and too little probability to central values.

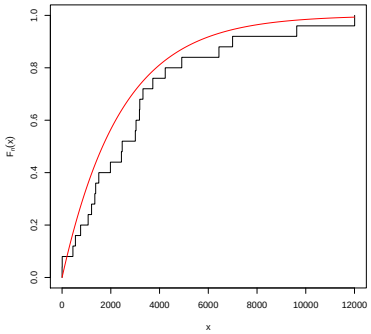
(iv) The exponential distribution assigns too much probability to central values and too little probability to tail values.

(c) Which of the following plots shows the empirical distribution function? Justify your answer.

(i)



(ii)



(iii)