ACSC/STAT 4703, Actuarial Models II
Fall 2015
Toby Kenney
Homework Sheet 5
Due: Friday 20th November: 10:30 PM

Basic Questions

1. An insurance company is modelling claim data as following a Weibull distribution with $\tau = 3$. It collects the following sample of claims:

```
1,445  1,392  1,655  1,260  1,525  1,604  3,134  2,095  1,447
1,304  1,350  1,793  1,945  888  1,485  1,998  1,916  2,774
1,482  705  1,427  705  3,024
```

   The MLE for θ is

$$\left(\frac{\sum_{i=1}^{23} {X_i}^3}{23}\right)^{\frac{1}{3}} = 1891.194$$

   Graphically compare this empirical distribution with the best fitting Weibull distribution with $\tau = 3$. 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 $F^*(x)$.

2. For the data in Question 1, calculate the following test statistics for the goodness of fit of the Weibull distribution with $\tau = 3$ and $\theta = 1891.194$ using:

   (a) The Kolmogorov-Smirnov test.

   (b) The Anderson-Darling test.

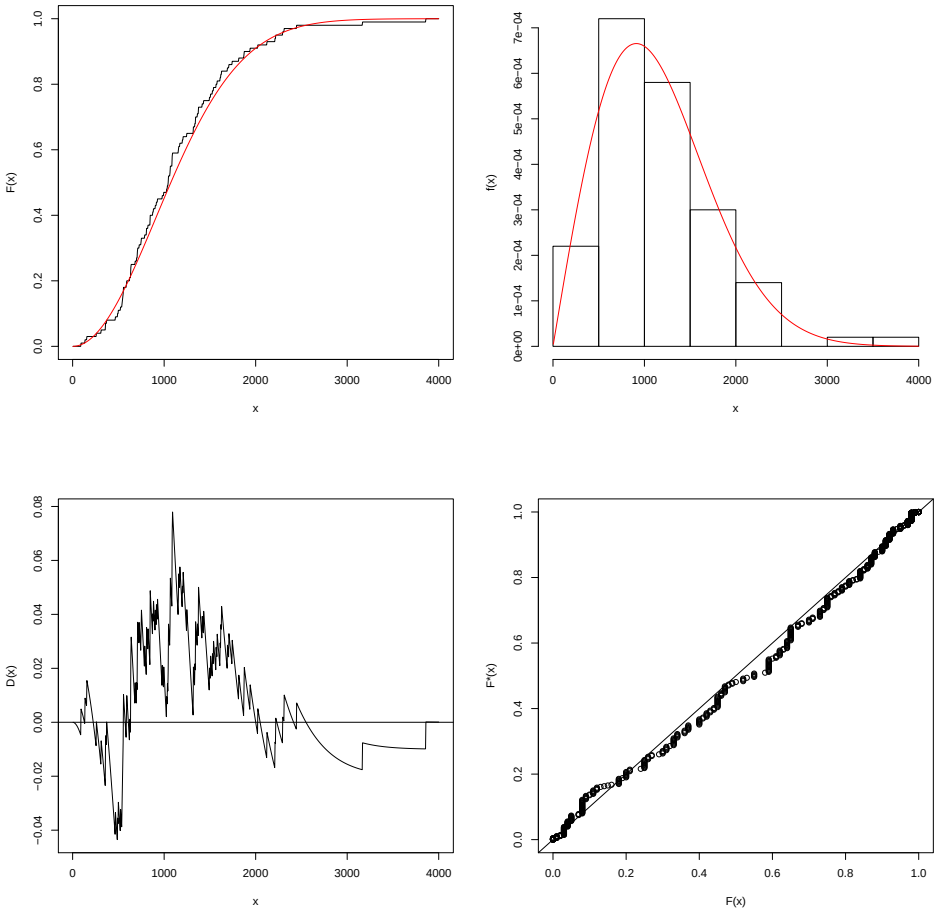
   (c) The chi-square test, dividing into the intervals 0–1500, 1500–2000, and more than 2000.

3. For the data in Question 1, perform a likelihood ratio test to determine whether a Weibull distribution with fixed $\tau = 3$, or a Weibull distribution with τ freely estimated is a better fit for the data. [The MLE for the general Weibull distribution is $\tau = 2.3831$ and $\theta = 1785.085$.]

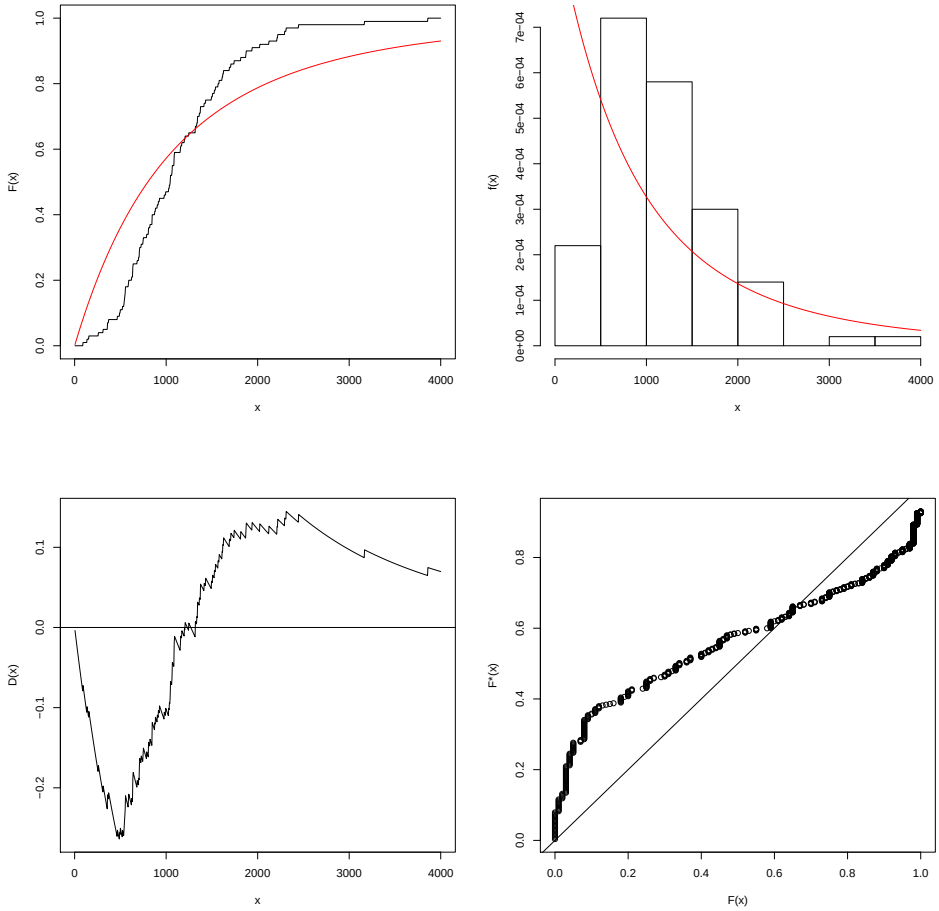
Standard Questions

4. An insurance company is modelling a data set. It is considering 3 models, each with 1 parameter to be estimated. Below are various diagnostic plots of the fit of each model.

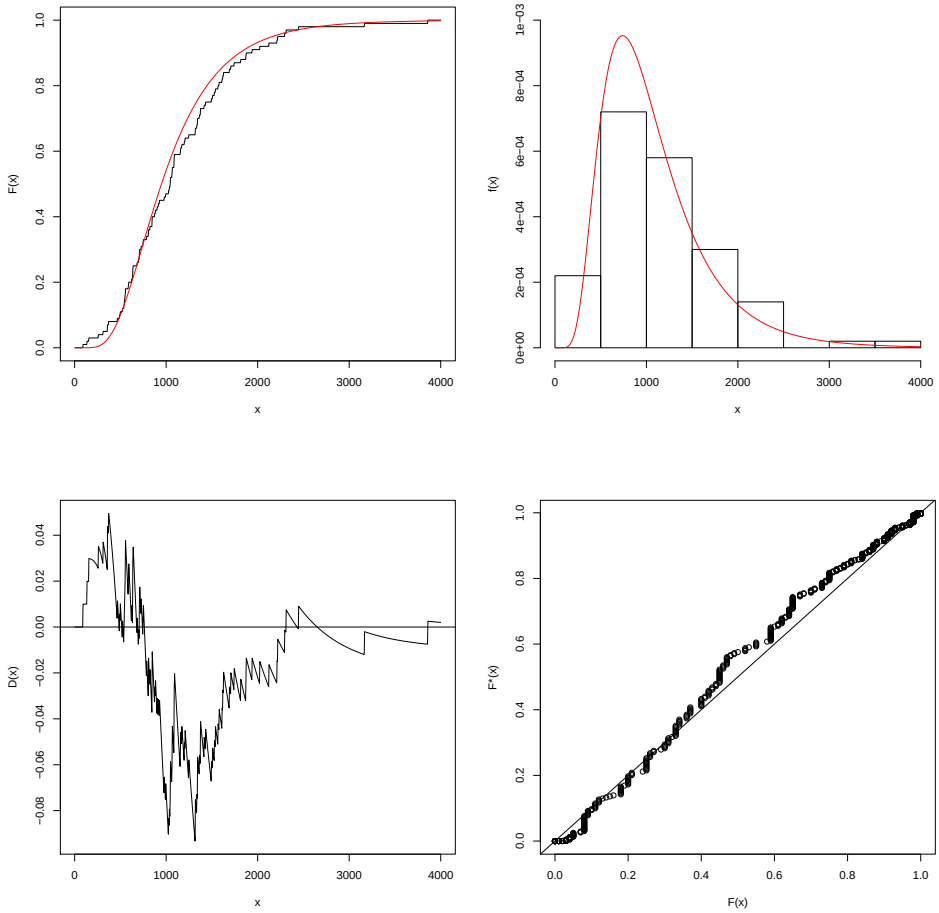
Model I

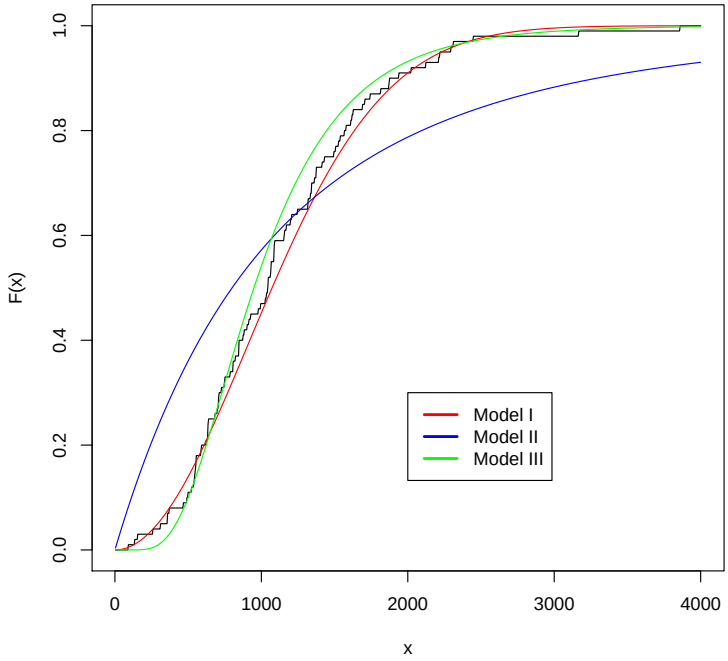


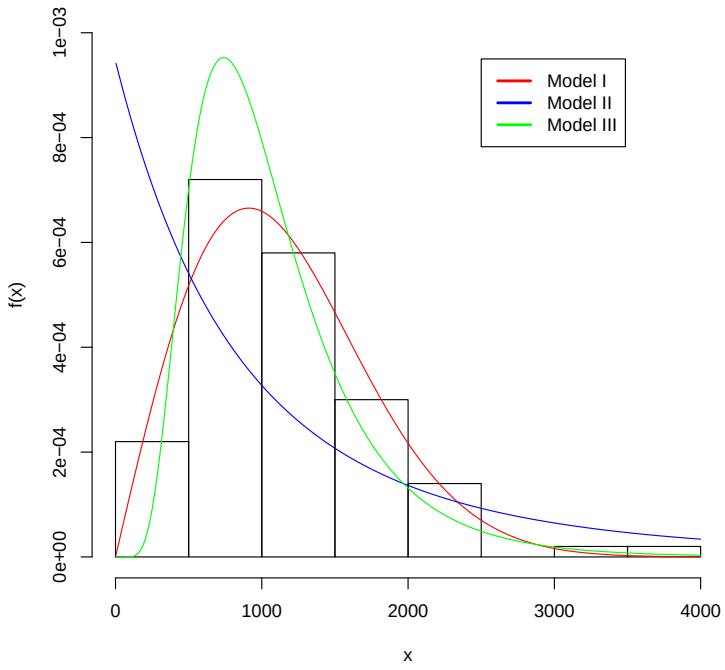
Model II



Model III




F(x)
0.0
0.2
0.4
0.6
0.8
1.0
0
1000
2000
3000
4000
x
Model I
Model II
Model III



Determine which model they should use for the data in the following situations. Justify your answers.

(a) Which model should they choose if accurately estimating (right-hand) tail probabilities is most important, and it is particularly important not to underestimate tail probabilities?

(b) The company is considering imposing a deductible, and therefore wants to model the distribution very accurately on small values of x.

(c) The company uses the Kolmogorov-Smirnov statistic to decide the best model.