EXAM C QUESTIONS OF THE WEEK

S. Broverman, 2005

Week of December 5

You are given the following random sample of 20 losses from the random variable X:
20 , 22 , 31 , 38 , 42 , c , 57 , 58 , 61 , 67 , 68 , 72 , 72 , 77 , 83 , 84 , 85 , 91 , 96 , 97

You are given that $42 < c < 57$.

The smoothed empirical estimate of the 25-th percentile is found, say A.

The data is now grouped into the following intervals:
$(0\,,\,20]$, $(20\,,\,40]$, $(40\,,\,60]$, $(60\,,\,80]$, $(80\,,\,100]$.
The empirical estimate of the cdf of X is also found based on the ogive of the grouped data set, and the estimate of the 25-th percentile is B.

Find the value of c for which $A = B$.

Solution can be found below.

Week of December 5 - Solution

With $n = 20$ data points, the smoothed percentiles of the given data points are

x	Smoothed Percentile
20	$\frac{1}{21} = 0.047619$
22	$\frac{2}{21} = 0.095238$
$\vdots$	$\vdots$
42	$\frac{5}{21} = 0.238095$
c	$\frac{6}{21} = 0.285714$
$\vdots$	$\vdots$

Since 0.25 is between $\frac{5}{21} = 0.238095$ and $\frac{6}{21} = 0.285714$, the smoothed empirical estimate of the 25th percentile will be between $x = 42$ and $x = c$. The smoothed empirical 25th percentile A is found by linear interpolation. The proportion of the way that A is between $x = 42$ and $x = c$ is the same proportion of the way that 0.25 is between $\frac{5}{21} = 0.238095$ and $\frac{6}{21} = 0.285714$, which is the same proportion of the way that 5.25 is between 5 and 6 (we get these by multiplying 0.25, 0.238095 and 0.285714 by 21). We see that 5.25 is $\frac{1}{4}$ of the way from 5 to 6, and therefore A is $\frac{1}{4}$ of the way from $x = 42$ to $x = c$. Therefore, $A = 42 + \frac{c-42}{4}$ is the smoothed empirical estimate of the 25th percentile.

Based on the grouping of the data, there is 1 observation in $(0, 20]$ so the empirical estimate of the cdf at $x = 20$ is $\frac{1}{20} = .05$. There are 3 observations in the interval $(20, 40]$, for a total of 4 observations in the interval $(0, 40]$, so the empirical estimate of the cdf at $x = 40$ is $\frac{4}{20} = .20$. There are 4 observations in the interval $(40, 60]$, for a total of 8 observations in the interval $(0, 60]$, so the empirical estimate of the cdf at $x = 60$ is $\frac{8}{20} = .40$.
The ogive is based on interpolating between interval endpoints. The empirical estimate of the 25-th percentile B is the linearly interpolated point between 40 and 60, since the empirical estimate of $F(40)$ is .20 and the empirical estimate of $F(50)$ is .4. Therefore, since .25 is $\frac{1}{4}$ of the way from .20 to .40, B must be $\frac{1}{4}$ of the way from 40 to 60, so $B = 45$.

In order to have $B = A$, we must have $42 + \frac{c-42}{4} = 45$, so that $c = 54$.