smfprob6 23.2.2017

PROBLEMS 6: NON-PARAMETRICS

Q1. *Brownian bridge covariance.*

Show that Brownian bridge B_0, where

$$B_0(t) := B_t - tB_1 \qquad (t \in [0,1]),$$

is

$$cov(B_0(s), B_0(t)) = \min(s,t) - st \qquad (s,t \in [0,1]).$$

Q2. *Median and breakdown point.*

The *breakdown point* of a statistical estimator is the proportion of the sample that can go off to infinity without dragging the estimator off with it. Thus the *mean* has breakdown point zero: for, if a single point in a finite sample goes off to $\pm\infty$, it drags the sample mean off to $\pm\infty$ with it.

The *sample median* is the middle reading (if there is a unique one – e.g., if the sample size is odd and there are no repetitions) – the 'score of the middle-ranking candidate', in examining terms. Show that the median has breakdown point $\frac{1}{2}$.

Note.

1. There is a whole subject here, of *robustness*. See for example

Frank R. HAMPEL, A general qualitative definition of robustness. *Annals of Math. Statistics* **42**:6 (1971), 1887-1896,

Peter J. HUBER & Elvezio M. RONCHETTI, *Robust statistics*, 2nd ed., Wiley, 2009 (1st ed., Huber, 1981).

2. With the population median med defined as the point with half the probability to the left/right (with appropriate change if there is an atom at the 'cross-over point'), and the sample median Med defined as above, we quote that, when there is a suitably smooth population density f, the median is asymptotically normal:

$$Med \sim N(med, \frac{1}{4nf(med)^2}).$$

Q3. *Semi-interquartile range (SIQ).*

The upper quartile $Q_{3/4}$ is the point with a quarter of the readings above it; the lower quartile $Q_{1/4}$ the point with a quarter of the readings below it; the semi-interquartile range is

$$SIQ := \frac{1}{2}(Q_{3/4} - Q_{1/4}).$$

Show that the quartiles have breakdown point 1/4, and SIQ has breakdown point 1/2.

Note.

1. While the standard deviation (SD) σ is not always defined, SIQ is. When both exist,

$$SIQ \leq \sqrt{2}\sigma,$$

and $\sqrt{2}$ here is best-possible. See e.g.

A. RÉNYI, *Wahrscheinlichkeitsrechnung* (1970), IV.15.

2. Thus SIQ gives a robust and more general alternative to the SD as a *measure of dispersion*.

NHB