STATISTICS FOR BUSINESS

TEACHER: JOSEMARI SARASOLA

SPECIAL COVID-19

VALIDATION: RUNS TEST

By means of the runs test we test if sample is really random. If fact, data randomness is an usual condition to develop inference methods.

Examples: data in the order thet have been collected:

5 7 8 1 1 6 3 2 2 9 7 8 8 9 2 0 11 12 13

Were data collected randomly? Alfa: %10.

Calculate median. To do that, order data:

```
0  1  1  2  2  2  3  5  6  7  7  8  8  8  9  9 11 12 13
```

Median is 7 (mid value)

Classify data according to relative position to median (below, -, above, +). If equal, remove data: -+------++++--+++

Plus 8, minus 9

Runs (sequences with equal sign): 6

Critical values for 8,9: 5-14 (look into tables).

6 is in the interval, so Ho: randomness accept. Data may be taken as random

Remark A: if data are dichotomous (men, women, for example; or ill/ not ill), no need to calculate medina, count sequences directly.

Remark B: if n,m (+,-) number is very big, tables are unuseful- We use a normal approximation: no of runs distributes normal (look slide no. 16). N1, n2 are n,m (no. of +,-). Then if alfa=%20 for example, calculate in the corresponding normal distribution values (z scores) than leave below and above it 10% (you know how to do that, if not take a look at problem no.120, d section). In no. of runs in the interval given by such values, accept randomness; otherwise, reject.

Remark C: if table value is missing (-), that means than on that side it's not possible to have a value rejecting Ho

TASKS

Solve problems 143, 144 and 145.