Statistics **NYU - Fall 1997**
Prof. Giorgio Topa **V31.0018 - 001**

ANSWERS TO MIDTERM EXAM

Question 1.

46	•
47	
48	•
49	
50	•
51	• •
52	• • •
53	• • • •
54	• • • • • •
55	• • • • • • •
56	• • • •
57	• • •
58	• • • •
59	• • •
60	• •
61	•
62	
63	• •
64	•
65	•
66	
67	• •
68	
69	•
70	•

The fastest way to solve this problem is to make the dot-plot on the left and use it to calculate the measures of central tendency asked in (b).

(a) See dot-plot on the left

(b) Mean: $(46+48+50+2\times51+3\times52+\ldots)/50=56.7$
Median: average of 25th and 26th observation. That is, it equals 55.5
Mode: Most frequent observation, i.e. 55.

(c) Median < Mean. Hence, distribution is right skewed.

Question 2.

Since there are 25 observations in the sample the first quartile of this sample is somewhere in between the 6th and 7th ordered observation. That is, it will be somewhere in between 65 and 66. Because we are calculating the first quartile, I have chosen for $Q_1=0.75\times65+0.25\times66=65.25$.

Question 3.

Given: The minutes of advertising per hour of broadcast time is a random variable with $\mu=3$ and $\sigma=2.1$.
Asked: Is an outcome of 7 a likely realization of this random variable.
Answer: If we calculate the z-score associated with 7 we obtain $z=(7-3)/2.1\approx1.90$. Hence $z<2$ and thus 7 is in the 75% range according to Chebyshev's rule. Thus, on the basis of this one observation, the radio station's claim seems to be valid.

Question 4.

(a) *Asked*: Proportion of managers that had earned at least one college degree.
Answer: This proportion equals
(# managers with at least one college degree)/(# managers)
which is *(76+48)/160=0.775*.

(b) *Asked*: $P(\text{advanced degree} \cap \text{good manager})$
Answer: $P(\text{advanced degree} \cap \text{good manager})=$
(# good managers with advanced degree)/(# managers)=12/160=0.075

(c) *Asked*: $P(\text{no college background} \mid \text{fair manager})$
Answer: $P(\text{no college background} \mid \text{fair manager}) =$
(# fair managers with no college background)/(# fair managers)=5/87=0.058

(d) *Asked*: $P(\text{advanced degree} \cup \text{good manager})$
Answer: $P(\text{advanced degree} \cup \text{good manager})=P(\text{advanced degree}) + P(\text{good manager}) - P(\text{advanced degree} \cap \text{good manager})=39/160+48/160-12/160=0.469$

(e) Such a frequency table can be reproduced in SAS with the following syntax:

```
proc freq;
tables degree*rating;
run;
```

where *degree* and *rating* are the variables that correspond to the data on the respective topics.

Question 5.

Given:

- *30* cameras
- 8 of which are defective
- 2 of the 30 are drawn at random.
- Shipment not accepted when a defective camera is drawn.

Define:

- D_i : { i^{th} camera drawn is defective } *for i=1,2*
- x : { # defective cameras drawn }. $x=0,1,2$.

(a) *Asked: P(x=2)*
Answer:
$P(x=2)=P(D_1 \cap D_2)=P(D_1)P(D_2 \mid D_1)=(8/30)(7/29)=0.064$

(b) *Asked*: $P(x \geq 1)$
Answer:
$P(x \geq 1)=1-P(x=0)=1- P(D'_1 \cap D'_2)=1-P(D'_1)P(D'_2 \mid D'_1)=1-(22/30)(21/29)=0.469$

Question 6.

There are two valid interpretations of this problem. In the following both are given.

Interpretation 1: Independence of probability of failure.
Given:

- $p=4/20=0.20$ is probability of transplant failure. (in terms of binomial: "success")
- $n=3$ is number of patients checked.
- All patients are <u>independent</u>.

Define:

- x : { *number of patients out of n=3 for which transplant fails* }
- x has binomial distribution with $n=3$ and $p=0.20$.

(a) *Asked*: $P(x=3)$

Answer: $p(x=3)=\binom{n}{3}p^3(1-p)^{n-3}=\binom{3}{3}0.2^3 0.8^0=0.2^3=0.008$

(b) *Asked*: $P(x\geq 1)$

Answer: $p(x\geq 1)=1-p(x=0)=1-\binom{n}{0}p^0(1-p)^{n-0}=1-0.8^3=0.488$

(c) See the definition above.

Interpretation 2: Dependence of probability of failure.
Given:

- 20 patients
- 4 of these 20 have transplants that will fail.
- 3 of these 20 are drawn at random.

Define:

- F_i : { i^{th} patient drawn has failing transplant } *for* $i=1,2,3$.
- x : { # patients with failing transplant drawn }, where $x=0,1,2,3$.

(a) *Asked*: $P(x=3)$

Answer: This is similar to question 5. That is
$P(x=3)=P(F_1\cap F_2\cap F_3)=P(F_1)P(F_2/F_1)P(F_3/F_1\cap F_2)=(4/20)(3/19)(2/18)=0.004$

(b) *Asked*: $P(x\geq 1)$

Answer: $P(x\geq 1)=1-P(x=0)=1-P(F'_1\cap F'_2\cap F'_3)=$
$1-P(F'_1)P(F'_2/F'_1)P(F'_3/F'_1\cap F'_2)=1-(16/20)(15/19)(14/18)=0.509$

Question 7.

Given:

- $p=0.90$ is probability that student is excited about statistics.
- $n=20$ students randomly selected.
- *assume*: students that are selected are indepent. Hence, each have probability of 90% of being excited about stats.

Define:

- x : { # excited students out of $n=20$ that are selected }
- x has binomial distribution with $n=20$ and $p=0.9$.

(a) *Asked*: expectation of x, i.e. μ.
Answer: $\mu=np=20\times0.9=18$.

(b) *Asked*: variance of x, i.e. σ^2.
Answer: $\sigma^2=np(1-p)=20\times0.9\times0.1=1.8$.

Question 8.

Given:

- 25 welfare workers
- 6 of these 25 have been giving illegal deductions.
- 2 workers selected at random.

Define

- *x : { # workers selected that have committed fraud }*
- F_i : { i^{th} worker that is selected is fraudulent }, $i=1,2$

Asked: $P(x=0)$

Answer: $P(x=0)=P(F'_1 \cap F'_2)=1- P(F'_1)P(F'_2 / F'_1)=(19/25)(18/24)=0.570$