MATH 333A: Probability & Statistics. $\boxed{\text{Examination \#1}}$ (Fall 2008)

October 15, 2008 NJIT **Solutions**

Score	
#1	
# 2	
#3	
#4	
#5	
Total	

Name:	Student ID:	Section

Instructors: R. Dios, A. Jain, and J. Porus

➔ **Must show all steps for each problem to receive full credit.**

I pledge my honor that I have abided by the Honor System. ____________________________
(Signature)

1. Richard Brown and Gretchen Davis have analyzed the following data to compare the ages of actors and actresses at the time they won Oscars:

Actors:	32 37 36 51 53 33 61 35 45 55 39 76 37 42 40 32 60 38 56 48 48 40 43 62 43 42 44 41 56 39 46 31 47 45 60 46 40 36
Actresses:	50 44 35 80 26 28 41 21 61 38 49 33 74 30 33 41 31 35 41 42 37 26 34 34 35 26 61 60 34 24 30 37 31 27 39 34 26 25 33

(a) Construct a back-to-back stem-and-leaf plot for the above data. (7 pts)
(b) Find the median ages. (7 pts)
(c) Comment on the shapes of the distributions of the ages of actors and actresses and any similarities/differences between the two distributions. (6 pts)

1 (a) To construct a back-to-back stem-and-leaf plot for the given data.

$N = 39$ for Actresses $N = 38$ for Actors

Leaf unit = 1.0

Stem: Tens digit
Leaf: ones digit

Actors	Stem	Actresses
	2	145666678
998776653221	3	0011333444455557789
8876655433221000	4	111249
66531	5	0
2100	6	011
6	7	4
	8	0

(b) To find the Median ages

(i) Actors :

Index for locating the median : $2\left(\frac{n+1}{4}\right)$

$n = 38$

$$\frac{38+1}{2} = \frac{39}{2} = 19.5$$

Median would be average of observations at location 19 & 20 in ranked data

$$\text{Median} = \frac{43+43}{2} = 43$$

$$\boxed{\text{Median age for Actors} = 43}$$

ii) Actresses:

Index for locating the median = $\frac{2(n+1)}{4}$

$n = 39$

$\frac{39+1}{2}$ = 20^{th} observation in ranked data

$\boxed{\text{Median age for actresses} = 34}$

c) The distributions of ages of actors and actresses are each right skewed.
The distribution of the ages of actors is less variable than the distribution of the ages of the actresses

2. The amounts of contamination (ppm) in a sample of 27 plastic specimens are given below:

30	30	60	63	70	79	87	90	101
102	115	118	119	119	120	125	140	145
172	182	183	191	222	244	291	350	511

(a) Compute the sample mean. (5 pts)
(b) Compute the median and quartiles. (5 pts)
(c) Construct a box-plot. (5 pts)
(d) Comment on your box plot of contamination. What are the important features of this plot? (5 pts)

2) a) To compute the sample mean $\bar{x}$

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{4059}{27} = 150.33$$

b) To compute the median & quartiles

Index for locating $Q_1 = \frac{n+1}{4} = 7$

Index for locating Q_2 or median $= \frac{2(n+1)}{4} = 14$

Index for locating $Q_3 = \frac{3(n+1)}{4} = 21$

$n = 27$

Q_1 is the 7^{th} observation in ranked data

$Q_1 = 87$

Q_2 or median is the 14^{th} observation in ranked data

Q_2 or Median $= 119$

Q_3 is the 21^{st} observation in ranked data

$Q_3 = 183$

c) To construct a box plot

Minimum $= 30$

Maximum $= 511$

$Q_1 = 87$

$Q_2 = 119$

$Q_3 = 183$

$IQR = Q_3 - Q_1 = 183 - 87 = 96$

$1.5(IQR) = 1.5(96) = 144$

$3(IQR) = 3(96) = 288$

$Q_1 - 1.5\ IQR = 87 - 144 = -57$

$Q_1 - 3\ IQR = 87 - 288 = -201$

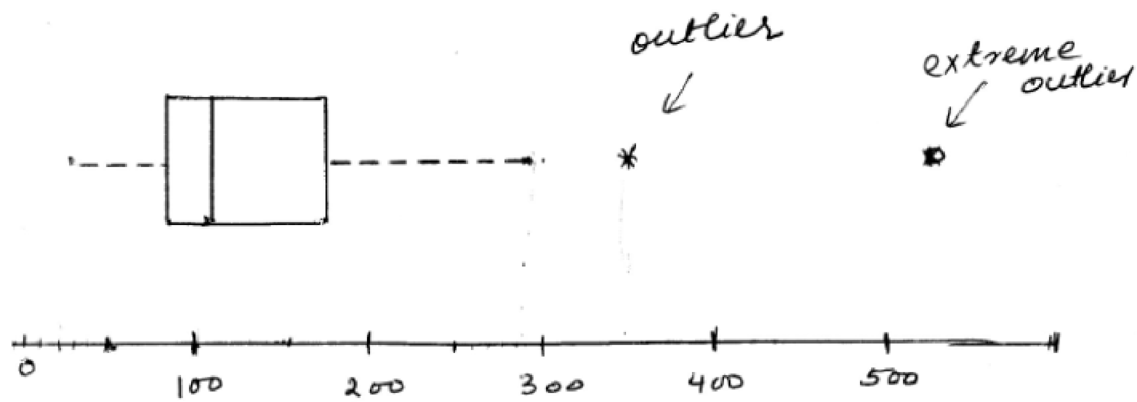
$Q_3 + 1.5\ IQR = 183 + 144 = 327$

$Q_3 + 3\ IQR = 183 + 288 = 471$

Data point 511 lies beyond 471, so would be an extreme outlier

Data point 350 lies beyond 327 but below 471; so would be an outlier.

Whiskers would extend to 30 & 291



d) Shape of the distribution: Skewed Right
There is an outlier at 350 and an extreme outlier at 511.

MATH 333 (A): Probability & Statistics. Examination #1 (Fall 2008)

3. Of patients going to a primary care physician's (PCP's) office, 30% are referred to a specialist, 40% require lab work, and 35% require neither lab work nor a referral to a specialist.

   (a) Determine the probability that a visit to a PCP's office results in both lab work and a referral to a specialist. (7 pts)
   (b) Consider the case when five records of patients' visits to a PCP's office are selected at random. Find the probability that all 5 of them are referred to a specialist. (7 pts)
   (c) Consider the case when another five records of patients' visits to a PCP's office are selected at random. Find the probability that at least one of them is referred to a specialist. (6 pts)

3) Let S denote the event "referred to a specialist"
Let L denote the event "require lab work"

$P(S) = 0.30$

$P(L) = 0.40$

$P(S' \cap L') = 0.35$

	L	L'	
S	0.05	0.25	0.30
S'	0.35	0.35	0.70
	0.40	0.60	1

a) Probability that a visit to a PCP's office results in both labwork and a referral to a specialist: $P(L \cap S) = 0.05$

b) Probability that one of them is referred to a specialist $P(S) = 0.30$

Assuming independent events, Probability that all five of them are referred to a specialist

$$P(S_1 S_2 S_3 S_4 S_5) = (0.30)(0.30)(0.30)(0.30)(0.30)$$
$$= 0.00243$$

c) Probability that at least one of them is referred to a specialist

= 1 − Probability that none are refered to a specialist

$= 1 - (0.70)^5$

$= 1 - 0.16807 = 0.83193$

4. A study of automobile accidents produced the following results:

Model Year	Proportion of all vehicles	Probability of involvement in an accident during the first half of 2008
2006	0.16	0.05
2007	0.18	0.02
2008	0.20	0.01
2005 or older	0.46	0.06

(a) An accident report for the first half of 2008 is chosen at random. What is the probability that the Model year for the car in this accident is 2008? (10 pts)

(b) A motor vehicle registration at the N.J. Motor Vehicle Bureau is selected at random. What is the probability that the model year of the vehicle in this registration record is 2005 or older and it is not involved in any accident during the first half of 2008? (10 pts)

4) To find:

(a) Probability that the model year for the car in this accident is 2008.

Let I denote "involvement in an accident during the first half of 2008"

Let A denote model year 2006

B denote model year 2007

C denote model year 2008

D denote model year 2005 or older

$P(A) = 0.16 \quad P(B) = 0.18 \quad P(C) = 0.20 \quad P(D) = 0.46$

$P(I|A) = 0.05 \quad P(I|B) = 0.02 \quad P(I|C) = 0.01$

$P(I|D) = 0.06$

To find $P(C|I)$

Using Bayes' theorem

$$P(C|I) = \frac{P(I|C)P(C)}{P(I|C)P(C) + P(I|A)P(A) + P(I|B)P(B) + P(I|D)P(D)}$$

$$= \frac{(0.01)(0.20)}{(0.01)(0.20) + (0.05)(0.16) + (0.02)(0.18) + (0.06)(0.46)}$$

$$\boxed{P(C|I) = 0.0485}$$

(b) To find $P(D \cap I')$

$$P(D \cap I') = P(D)P(I'|D) = (0.46)(1-0.06)$$
$$= (0.46)(0.94)$$
$$= 0.4324$$

5. A lot of 50 components contains 5 defective components. One of the 50 components is drawn at random on Day 1 and it is used in Assembly 1. A second component is drawn at random from the remaining 49 components on Day 2 and it is used in Assembly 2. Let F be the event that the first component drawn is defective, and S is the event that the second component drawn is defective. (5 pts each)

   (a) Find probability P(F).
   (b) Find probability P(F and S).
   (c) Find probability P(S).
   (d) Are events F and S independent? Explain why or why not.

5) a) $P(F) = \frac{5}{50} = 0.1$

b) $P(F \cap S) = P(F)\,P(S|F) = \left(\frac{5}{50}\right)\left(\frac{4}{49}\right) = 0.0082$

c) $P(S) = P(S|F)P(F) + P(S|F')P(F')$
$$= \left(\frac{4}{49}\right)\left(\frac{5}{50}\right) + \left(\frac{5}{49}\right)\left(\frac{45}{50}\right) = 0.1$$

d) $P(F|S) = \frac{P(S \cap F)}{P(S)} = \frac{0.0082}{0.1} = 0.082$

$P(F) = 0.1$

$P(F|S) \neq P(F)$

Hence, events F & S are not independent.