AMS 310 Survey of Probability and Statistics

Midterm Exam I

Name__________ Student ID__________

Total Score____________________

Each problem is 25 points. The last problem is for extra credit. Good Luck!
(1). The following are the scores from the final exam of a graduate course:

100, 95, 84, 62, 97, 82, 93, 32, 66, 78, 83, 85, 79, 82, 71, 89

(a). Find the mean, median, and standard deviation of the grade.

(b). Using the class interval (bin width) as 10, draw a Pareto chart and a dot-plot for the grade.

(c). Identify students who are above the third quartile.

Solution:

(a). Sort the data in descending order:

32, 62, 66, 71, 78, 79, 82, 82, 83, 84, 85, 89, 93, 95, 97, 100

The median is 82.5, the third quartile is 91, mean is 80, variance is 218.

(b). The top ten percentile are 95, 97, and 100.

(c).
Figure 2: Pareto Diagram
A random variable can have values of 0, 1, 2, 3, 4, 5, 6, 7.

(a). It is known that $P(X \leq 4) = 3P(X \geq 5)$, find the probability of $P(X \leq 4)$ and $P(X \geq 5)$.

(b). If $P(X \leq 6) = 0.9$ and $P(X \geq 2) = 0.8$, find the probability for $P(2 \leq X \leq 6)$.

(c). If $P(1 \leq X \leq 5) = 0.8$, $P(3 \leq X \leq 6) = 0.6$, and $P(3 \leq X \leq 5) = 0.5$, find the probability for $P(1 \leq X \leq 6)$.

Solution

(a). Let $P(A) = P(X \leq 4)$,

then $P(\bar{A}) = P(X \geq 5)$,

we have $P(A) + P(\bar{A}) = 3P(\bar{A}) + P(\bar{A}) = 1$, $4P(\bar{A}) = 1$,

hence $P(\bar{A}) = 0.25$, $P(A) = 0.75$.

(b). Let $P(B) = (X \leq 6)$,

$P(C) = P(X \geq 2)$,

$P(2 \leq X \leq 6) = P(B \cap C)$

Since $P(B \cup C)$ covers all values of the random variable, therefore

$P(B \cup C) = 1 = P(B) + P(C) - P(B \cap C)$

hence $P(B \cap C) = P(B) + P(C) - 1 = 0.9 + 0.8 - 1 = 0.7$

(c). Let $P(D) = P(1 \leq X \leq 5)$,

$P(E) = P(3 \leq X \leq 6)$,

we also have $P(D \cap E) = P(3 \leq X \leq 5)$,

hence

$P(1 \leq X \leq 6) = P(D \cup E) = P(D) + P(E) - P(D \cap E) = 0.8 + 0.6 - 0.5 = 0.9$. 
(3) Mechanical failure is accounted for 40 percent of airline accidents, the others are due to human operation errors. Find the probability that in the next ten airline accidents:

(a). None of them is due to human operation error.
(b). All of them are due to human operation error.
(c). Five of them are due to human operation error.

**Solution:** For each accident, it is a Bernoulli distribution with $p = 0.4$. For 10 accidents, it is binomial distribution.

(a). All 10 are due to mechanical failur, so the probability is

$$B(x = 10; n = 10, p = 0.4) = \binom{10}{10} 0.4^{10} = 0.4^{10}.$$

(b). None is due to mechanical failure,

$$B(x = 0; n = 10, p = 0.4) = \binom{10}{0} 0.6^{10} = 0.6^{10}.$$

(c). Five are due to mechanical failure,

$$B(x = 5; n = 10, p = 0.4) = \binom{10}{5} 0.4^5 0.6^5.$$
(4). (a). About 40 percent of women who take pregnancy test are actually carrying a baby. The testing is 99 percent accurate. If a woman takes a test and the result is positive (showing pregnancy), what is the probability that she is actually carrying a baby?

(b). Among all women who take mammography test each year, only 0.1 percent of them actually have breast cancer. Suppose the mammography is 99 percent accurate and a woman is tested positive for breast cancer, what is the probability she actually does not have cancer?

**Solution:**

(a). Let $A$ be the event of non-pregnant woman taking a test and $\overline{A}$ be the event of pregnant woman taking a test, we have $P(A) = 0.6$ and $P(\overline{A}) = 0.4$. Let $B$ denote the event that the test is positive, then from total probability, we have

\[
P(B) = P(B/A)P(A) + P(B/\overline{A})P(\overline{A}) = 0.01 \times 0.6 + 0.99 \times 0.4 = 0.402
\]

\[
P(\overline{A}/B) = P(\overline{A} \cap B)/P(B) = 0.99 \times 0.4/0.402 = 0.985
\]

The probability that a woman has positive result and actually carry a baby is 0.985.

The probability that a woman has positive result but does carry a baby is 0.015.

(b). Let $A$ be the event that a healthy woman taking a mammography test and $\overline{A}$ be the event that a woman with breast cancer taking a test, we have $P(A) = 0.999$ and $P(\overline{A}) = 0.001$. Let $B$ denote the event that mammography test is positive, then from total probability, we have

\[
P(B) = P(B/A)P(A) + P(B/\overline{A})P(\overline{A}) = 0.999 \times 0.01 + 0.001 \times 0.99 = 0.01098
\]

\[
P(A/B) = P(A \cap B)/P(B) = 0.999 \times 0.01/0.01098 = 0.91
\]

The probability that a woman has positive mammography result but is actually healthy is 0.91 (false positive).
(5). If a share of stock is allowed to go up and down 50 cent each day and if the probability of going up or down is 50 percent. Let the stock starts from 20 dollars per share, after 40 days, what is the probability that the stock price is below 15 dollars or above 25 dollars?

**Solution:** The stock will have the price $15 \leq S \leq 25$ if the number of ups is between $15 \leq x \leq 25$. The stock daily gain is a Bernoulli probability and for 40 days,

$$P(15 \leq x \leq 25) = \sum_{x=15}^{25} B(x; 20, 0.5) = \sum_{x=15}^{25} \binom{40}{x} 0.5^{40}$$

The probability of stock price beyond this range is

$$P(x < 15) + P(x > 25) = 1 - P(15 \leq x \leq 25) = 1 - \sum_{x=15}^{25} \binom{40}{x} 0.5^{40}$$
(6) (a). How many ways to make a 6 letter password without repetition?
(b). How many ways to make a 6 letter password allowing repetition?
(c). How many ways to form a committee with two women and two men from total of eight women and ten men.

Solution:

(a). There are 26 letters. To select 6 for password is a permutation

\[ N = 26P_6. \]

(b). When repetition is allowed, each letter in the password has 26 choices. Since each pick-up is independent of other pick-ups, therefore the number of ways is

\[ N = 26^6. \]

(c). The process can be decomposed into two stages: select two women from the eight women; and then select two men from the 10 men.

\[ N = \binom{8}{2} \binom{10}{2}. \]
(7). A random variable can take seven values: 0, 1, 2, 3, 4, 5, 6. The probability for each value is \( P(X = x) = C(x - 3)^2 \). Find the value of the parameter \( C \). Find the mathematical expectation (mean) and the variance of the random variable.

**Solution:** Since

\[
\sum_{x=0}^{6} P(X = x) = \sum_{x=0}^{6} C(x - 3)^2 = C(3^2 + 2^2 + 1^2 + 0 + 1^2 + 2^2 + 3^2) = 28C = 1,
\]

hence

\[ C = \frac{1}{28} \]

The mathematical expectation is

\[
\mu = E(X) = \frac{1}{28} \sum_{x=0}^{6} x(x - 3)^2
\]

The variance is

\[
\sigma^2 = E(X^2) - \mu^2 = \frac{1}{28} \sum_{x=0}^{6} x^2(x - 3)^2 - \left( \frac{1}{28} \sum_{x=0}^{6} x(x - 3)^2 \right)^2
\]
(8). (a). There are 6 white balls and 10 black balls in a bag. If a person randomly draws 4 balls from the bag, what is the probability that two balls are black?

(b). If the person put the ball back to the bag after each draw, what is the probability that in 4 draws, two are black?

(c). If the ratio of white and black balls is kept at 6/10, as the total number of balls in the bag increases, what is the best way to calculate the approximate probability of (a)?

**Solution:**

(a). This is hypergeometric probability distribution problem with \( N = 16 \), \( a = 10 \) and \( n = 4 \)

\[
P(x = 2) = G(x = 2; 40, 6, 4) = \binom{10}{2} \binom{6}{2} \binom{40}{4} \]

(b). This is a binomial probability distribution problem with \( n = 4 \) and \( p = 10/16 = 0.625 \)

\[
P(x = 2) = B(2; 4, 0.625) = \binom{4}{2} 0.625^2 0.375^2 \]

(c). When \( N \) is very large and \( a/N \) remain unchanged, the hypergeometric distribution can be approximated by the binomial distribution, therefore

\[
P(x = 2) \approx B(2; 4, 0.625)\]

should be used because the calculation is much easier.
Bon us problem (20 p oin ts)

(9). Pro v e:

\[ \sum_{k=0}^{x} \binom{n}{k} p^k (1 - p)^{n-k} + \sum_{k=0}^{n-x-1} \binom{n}{k} p^{n-k} (1 - p)^k = 1 \]

\[ \textbf{Proof:} \text{ Consider the second term in the LHS:} \]

\[ \sum_{k=0}^{n-x-1} \binom{n}{k} p^{n-k} (1 - p)^k = \sum_{k=0}^{n-(x+1)} \binom{n}{n-k} p^{n-k} (1 - p)^k. \]

For each term, change the index \( k' = n - k \), we have

\[ \sum_{k'=n}^{x+1} \binom{n}{k'} p^{k'} (1 - p)^{n-k'} = \sum_{k'=x+1}^{n} \binom{n}{k'} p^{k'} (1 - p)^{n-k'} = \sum_{k=x+1}^{n} \binom{n}{k} p^k (1 - p)^{n-k}. \]

Substitute back to the original LHS, we have

\[ LHS = \sum_{k=0}^{x} \binom{n}{k} p^k (1 - p)^{n-k} + \sum_{k=x+1}^{n} \binom{n}{k} p^k (1 - p)^{n-k} \]

\[ = \sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k}. = (p + (1 - p))^n = 1^n = 1 \]