Data Analysis and Decision Making Handout 1

Solutions of the in-class problems. The problems will be given from the Text book. Topic of this work includes the Exercises Section of the chapters 2 and 3.

Exercise 2.1

$P(X \geq 5) = 1 - P($

A four-sided die is engraved with the numbers 1 through 4 on its four different sides. Suppose that when rolled, each side (and hence each number) has an equal probability of being the bottom face when it lands. We roll two such dice. Let *X* be the sum of the numbers on the bottom faces of the two dice.
(a) What is the probability that *X* is at least five?
(b) How does your answer to (a) change if you are told that the bottom face of *the first die* has the number "3" on it?
(c) How does your answer to (a) change if you are told that the bottom face of *one of the dice* has the number "3" on it?

(a) $P(X \geq 5) = \frac{10}{16}$.

(b) $P(X \geq 5 \mid \text{First} = 3) = \frac{3}{16} / \frac{4}{16} = \frac{3}{4}$

(3) $P(X \geq 5 \mid \text{at least 3 in One die}) = \frac{5}{16} / \frac{7}{16}$.

$P(A|B) = \frac{P(A \cap B)}{P(B)}$

Exercise 2.7

It is a relatively rare event that a new television show becomes a long-term success. A new television show that is introduced during the regular season has a 10% chance of becoming a long-term success. A new television show that is introduced as a mid-season replacement has only a 5% chance of becoming a long-term success. Approximately 60% of all new television shows are introduced during the regular season. What is the probability that a randomly selected new television show will become a long-term success?

$0.6 \times 0.1 + 0.4 \times 0.05 = 0.08$

S_1: regular Season $P(S_1) = 0.6$ A: long terms

S_2: middle Season $P(S_2) = 0.4$

$P(A|S_1) = 0.1$ $P(A|S_2) = 0.05$

0.6
0.4

?

Exercise 2.10

A hardware store has received two shipments of halogen lamps. The first shipment contains 100 lamps, 4% of which are defective. The second shipment contains 50 lamps, 6% of which are defective. Suppose that Emanuel picks a lamp (at random) off of the shelf and purchases it, and he later discovers that the lamp he purchased is defective. Is the defective lamp more likely to come from the first shipment or from the second shipment?

~~Assume that~~ $\begin{cases} X_1: \text{defective lamp from first shipment} \\ X_2: \text{defective lamp from second shipment} \end{cases}$ S_1: first, S_2: Second

$P(X_1) = \frac{100 \times 4\%}{\cancel{100+50}} = \frac{4}{150}$

$P(X_2) = \frac{50 \times 6\%}{150} = \frac{3}{150}$

$P(D|S_1) = 0.04$; $P(D|S_2) = 0.06$

$P(S_1|D) = \frac{P(S_1) \cap P(D)}{P(D)} = \frac{P(D|S_1) \cdot P(S_1)}{P(D)}$; $P(S_1) = \frac{100}{150} = \frac{2}{3}$

$P(D) = P(D|S_1) \cdot P(S_1) + P(D|S_2) \cdot P(S_2) = 0.04 \times \frac{2}{3} + 0.06 \times \frac{1}{3} = \frac{7}{150}$

$P(S_1|D) = \frac{P(D|S_1) \cdot P(S_1)}{P(D)} = \frac{0.04 \cdot \frac{1}{3}}{\frac{7}{150}} = \frac{8}{14}$

$P(S_2|D) = \frac{P(D|S_2) \cdot P(S_2)}{P(D)} = \frac{0.06 \cdot \frac{1}{3}}{\frac{7}{150}} = \frac{6}{14}$

Exercise 2.20

A manufacturing company of children's clothing has three production plants located in Andover, Bedford, and Concord. The number of items produced per day (in 1,000s) in Andover has a mean of 91 and a standard deviation of 3.2; the mean daily production rate in Bedford is 67 with a standard deviation of 2.2; and the mean daily production rate in Concord is 69 with a standard deviation of 5.7. Let X be the total number of items produced per day by the company at all three sites (in 1,000s).

(a) What is the mean of X?

(b) Suppose that production levels in Andover, Bedford, and Concord are independent. What is the variance of X? What is the standard deviation of X?

(a). $E(X) = E(A) + E(B) + E(C) = 91 + 67 + 69 = 227$

(b) $Var(X) = Var(A) + Var(B) + Var(C) = (3.2)^2 + (2.2)^2 + (5.7)^2 = 47.57$

$SD(X) = \sqrt{47.57} = 6.9$

3 decimal point in exam

扫描全能王 创建

Data Analysis and Decision Making Handout 2

Solutions of the in-class problems. The problems will be given from the Text book. Topic of this work includes the Exercises Section of the chapters 2 and 3.

Exercise 3.6

Winter lasts from December 21 through March 21. The average winter temperature in Boston is Normally distributed with mean μ = 32.5° F and standard deviation s = 1.59° F. In New York City, the average winter temperature is Normally distributed with mean μ = 35.4° F and standard deviation s = 2.05° F.

(a) What is the probability that the average winter temperature in Boston this coming winter will be above freezing (32° F)? $P(X>32)=1-P(X<32)=1-P(Z<\frac{32-32.5}{1.59})=1-P(Z<0.3)=1-0.618$

(b) Assume that average winter temperatures in Boston and New York are independent. What is the probability that the average temperature in Boston in the coming winter will be higher than in New York?

(c) Do you think the independence assumption above is reasonable?

(b) $D=Y-X<0$ $P(D<0)=P(Z<\frac{0-2.9}{2.59434})=0.1318224$

$\mu_D=2.9$

$\sigma_D=\sqrt{\sigma_x^2+\sigma_y^2}=\sqrt{1.59^2+2.05^2}=2.59434$

② $\hat{P}=\frac{X}{n}\sim N(P=0.522, \sqrt{\frac{pq}{n}}=0.0499)$

Exercise 3.10

According to a recent census, 52.2% of Boston residents are female. Suppose a group of 100 Bostonians is selected at random.

(a) What is the mean and the standard deviation of the number of female members of the group? ① $\mu=np=52.2$ & $\sigma=\sqrt{npq}=\sqrt{24.95}=5 \Rightarrow X\sim N(52.2, \sqrt{24.95})$

(b) Use the Normal approximation to find the probability that less than one half of the members of the group are female. $P(\hat{P}<0.5)=P(Z<\frac{0.5-0.522}{\sqrt{\frac{0.522\times0.48}{100}}})=P(Z<-0.44)=33\%$

(c) Use the Normal approximation to find the probability that the number of female members of the group is between 45 and 55.

① $P(45<\bar{X}<55)=P(\frac{0.45-0.52}{\sqrt{\frac{0.48\times0.52}{100}}}<\hat{P}<\frac{0.55-0.52}{\sqrt{\frac{0.48\times0.52}{100}}})=64\%$

② $P(45<\bar{X}<55)=P(Z<\frac{55-100\cdot0.52}{\sqrt{100\times0.48\times0.52}})-P(Z<\frac{45-100\cdot0.52}{\sqrt{100\times0.48\times0.52}})$

$=P(Z<0.6)-P(-1.40112)$

$=64\%$

Exercise 3.19 $\begin{cases} E(A)=0.1 \\ \sigma_A=0.04 \end{cases}$ $\begin{cases} E(B)=0.2 \\ \sigma_B=0.1 \end{cases}$ CORR$=0.2$

Investor's Market Times is a magazine that rates stocks and mutual funds and publishes predictions of stock performance. Suppose that *Investor's Market Times* has predicted that stock A will have an expected annual return of 10% with a standard deviation of 4%,
and stock B will have an expected annual return of 20% with a standard deviation of 10%. *Investor's Market Times* has also estimated that the correlation between the return of these stocks is0.20.

(a) What fraction of your portfolio should you invest in each of stocks A and B so that the expected annual return is 13%?

$0.1X + 0.2(1-X) = 0.13 \quad A=70\% \quad B=30\%$

(b) What is the standard deviation of the return of the portfolio that has an expected annual return of 13% (part (a))?

$S_0 = \sqrt{0.7^2 \times (0.04)^2 + 0.3^2 \times (0.1)^2 + 2 \times 0.7 \times 0.3 \times 0.04 \times 0.1 \times 0.2} = 0.0449 = 4.49\%$

(c) The magazine recommends 50 small international stocks as great buying opportunities in the next year. They claim that each of these 50 stocks has an expected annual return of 20% and a standard deviation of 20%. They also claim that these returns are independent. For diversification reasons, they recommend that investors invest 2% of their money in each stock and hold the portfolio for a year. What is the expected return and standard deviation of the annual return of this portfolio?

$\begin{cases} u=0.2 \\ sd=0.2 \end{cases}$

$E(\text{return}) = 50 \times 20\% \times 2\% = 0.2$

$\sigma(\text{return}) = \sqrt{50 \times (0.02)^2 \times (0.2)^2} = 0.02828$

Exercise 4.6 t-test

An investment analyst would like to estimate the average amount invested in a particular mutual fund by institutional investors. A random sample of 15 institutional investors' portfolios has been chosen. The observed sample mean of the amounts invested in the mutual fund by these 15 investors is $11.32 million and the observed sample standard deviation is $4.4 million. Assume the portfolios are distributed normally.

$0.9 = 1-\alpha \Rightarrow \frac{\alpha}{2} = 0.05$

(a) Construct a 90% confidence interval for the mean amount invested by all institutional investors in the mutual fund.

$11.32 \pm t_{0.05,14} \frac{4.4}{\sqrt{15}} = 11.32 \pm 1.761 \frac{4.4}{\sqrt{15}} \Rightarrow [9.31, 13.32]$

(b) Determine the required sample size in order to estimate the mean amount invested by all institutional investors in this mutual fund to within $500,000, at the 95% confidence level.

$Z_{0.025}\sqrt{\frac{\sigma^2}{n}} \le W \quad 1.96 \times \frac{S}{\sqrt{n}} < 0.5$

$\Rightarrow 1.96 \times \frac{4.4}{\sqrt{n}} < 0.5$

$n > \left(\frac{1.96 \times 4.4}{0.5}\right)^2$

$n > 298.$

Exercise 4.16

$1-\alpha=0.98 \quad \frac{\alpha}{2}=0.01 \quad Z_{0.01}=t_{0.01,\infty}=2.326$

An information technology service at a large university surveyed 200 students from the College of Engineering and 100 students from the College of Arts. Among the survey participants, 91 Engineering students and 73 Arts students owned a laptop computer. Construct a 98% confidence interval for the difference between the proportions of laptop owners among Engineering and Arts students.

$$\left(\frac{91}{200}-\frac{73}{100}\right) \pm Z_{0.01}\sqrt{\frac{\left(\frac{91}{200}\right)\left(\frac{109}{200}\right)}{200}+\frac{\left(\frac{73}{100}\right)\left(\frac{27}{100}\right)}{100}}$$

$[-0.4068, -0.1432]$