SOLUTION In-Class Assignment # 4 & 5 04.25.11

Answer the following questions on a separate sheet of paper.

1. The semi-annual returns for three stocks over an 18 year period are examined. The probability that Stock 1 will outperform the other two stocks is 50%, the probability that Stock 2 will outperform the other two stocks is 30%, and the probability that Stock 3 will outperform the other two stocks is 20%. Calculate the expected value **AND** variance of this distribution.

X	$f(x)$	$X * f(x)$	$X - E(x)$	$[X - E(x)]^2$	$[X - E(x)]^2 * f(x)$
1	.50	$1 * .50 = .50$	$1 - 1.7 = -.7$	$-.7^2 = .49$	$.49 * .50 = .2450$
2	.30	$2 * .30 = .60$	$2 - 1.7 = .3$	$.3^2 = .09$	$.09 * .30 = .0270$
3	.20	$3 * .20 = .60$	$3 - 1.7 = 1.3$	$1.3^2 = 1.69$	$1.69 * .20 = .3380$
		$E(x) = 1.7$			$\sigma^2 = .6100$

Expected Value:

$E(x) = \underline{1.7}$

Variance:

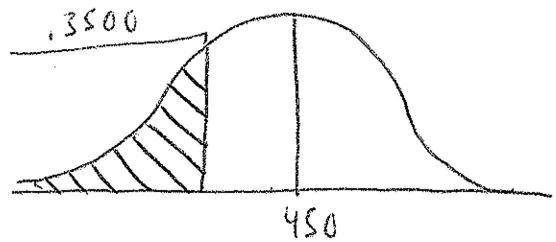
$\sigma^2 = \underline{.6100}$

2. The weekly demand for Baked Lay's potato chips at a certain Subway sandwich shop is a normally distributed random variable with a mean of 450 and a standard deviation of 80. What is the highest weekly demand for Baked Lay's potato chips for the bottom 35% of the weeks?

① $\bar{X} = 450$ $\sigma = 80$ $X = ?$

② $P(x < c) = .3500$

③



④ $z = -.39$

⑤ $-.39 = \dfrac{(x - 450)}{80}$

$-31.20 = X - 450$

$X = 418.80$

⑥ The highest demand level for Baked Lay's chips for the bottom 35% of the weeks is 418.80 bags.

3. Calculate the sample standard deviation **AND** the coefficient of variation for the following data set.

12 4 6 1 14

x	$(x-\bar{x})$	$(x-\bar{x})^2$
12	$12-7.4=4.6$	$4.6^2=21.16$
4	$4-7.4=-3.4$	$-3.4^2=11.56$
6	$6-7.4=-1.4$	$-1.4^2=1.96$
1	$1-7.4=-6.4$	$-6.4^2=40.96$
14	$14-7.4=6.6$	$6.6^2=43.56$
		119.2

$\bar{x}=7.4$

$$\theta=\sqrt{\frac{119.2}{(5-1)}}=\sqrt{29.8}=\underline{\underline{5.4589}}$$

$$CV=\frac{5.4589}{7.4}*100=\underline{\underline{73.77\%}}$$

4. A random sample of 36 statistics examinations was taken. The average score, in the sample, was 76.2 with a sample variance of 196. Calculate the 95% confidence interval for the true average examination score.

(1) $n = 36$ $\quad \bar{x} = 76.2$ $\quad \sigma^2 = 196$ $\quad CL = .95$

↳ convert to $\sigma = 14$

(2) $\alpha = 1 - .95 = .05$

$\alpha/2 = .025$

(3) $z_{.025} = 1.960$

(4) $76.2 \pm 1.960 \dfrac{14}{\sqrt{36}}$

76.2 ± 4.5733

71.6267 to 80.7733

(5) The true mean exam score will be between 71.6267 and 80.7733 based on a sample of 36 and a 95% confidence level.

5. A congressional committee wishes to estimate the average annual subsidy received by tenant farmers in a southern state. A random sample of 549 farms is taken, the sample mean is found to be $7,800 with a sample standard deviation of $750. Compute the 99% confidence interval estimate of the annual subsidy?

(1) $n = 549$ $\bar{x} = 7800$ $\sigma = 750$ $CL = .99$

(2) $\alpha = 1 - .99 = .01$

$\alpha/2 = .01/2 = .005$

(3) $z_{.005} = 2.576$

(4) $7800 \pm 2.576 \frac{750}{\sqrt{549}}$

7800 ± 82.4558

7717.5442 to 7882.4558

(5) At a 99% confidence level, the mean subsidy will be between \$7717.5442 to \$7882.4558 based on a sample of 549.

6. Business statistics students claim that they spend at least 3 hours studying for and working through statistics practice problems per week. A random sample of 34 college students enrolled in a business statistics course was asked how many hours he or she spent practicing working statistics problems. The sample average was found to be 2.4 hours per week with a standard deviation of 1.25 hours. Using the p-value method and $\alpha = .05$, determine if the students are studying the amount of time that was claimed.

① $\mu = 3$ $\quad n = 34$ $\quad \bar{x} = 2.4$ $\quad \sigma = 1.25$ $\quad \alpha = .05$

② $H_0: \mu \geq 3$

$H_A: \mu < 3$

③ p-value method

$$z = \frac{(2.4 - 3)}{1.25/\sqrt{34}} = \frac{-.6}{.2144} = -2.7985 \approx -2.80$$

④ p-value = .0026

⑤ Reject H_0 if p-value $\leq \alpha$

$.0026 \leq .05$ TRUE

⑥ Reject H_0: The data do not support the claim that students spend at least 3 hours studying statistics based on a sample of 34 and a 5% level of significance.

7. A highway patrol officer states that the average speed of cars traveling over a certain stretch of highway is more than 65 miles per hour which is faster than the posted speed of 55 miles per hour. A sample of 100 drivers is taken and the sample average is found to be 62.5 miles per hour with a standard deviation of 9.8 miles per hour. Using the confidence interval method and a level of significance of 0.10, draw a conclusion for this situation.

(1) $\mu = 65$ $\quad n = 100$ $\quad \bar{x} = 62.5$ $\quad \sigma = 9.8$ $\quad \alpha = .10$

(2) $H_0: \mu \geq 65$

$H_A: \mu < 65$

(3) Confidence interval method

$\alpha = .10$

$Z_{.10} = 1.282$

(4) $62.5 \pm 1.282 \dfrac{9.8}{\sqrt{100}}$

62.5 ± 1.2564

61.2436 to 63.7564

(5) Reject Ho if μ falls outside of the confidence interval

65 does fall outside of 61.2436 to 63.7564

(6) Reject H_0: The data does not support the statement that cars travel more than 65 miles per hour based on a sample of 100 and a ~~[illegible]~~ 10% level of significance.

8. A fast-food franchiser is considering building a restaurant at a certain location. Based on financial analyses, a site is acceptable only if the average number of pedestrians passing the location is 100 or more. The number of pedestrians observed for each of 40 hours was recorded and the average was found to be 103 with a standard deviation of 12. Can the fast-food franchiser conclude that the site is acceptable at a 0.01 level of significance? Use the critical value method to draw your conclusion.

(1) $\mu = 100$ $n = 40$ $\bar{x} = 103$ $\sigma = 12$ $\alpha = .01$

(2) $H_0: \mu \geq 100$

$H_A: \mu < 100$

(3) critical value method

$\alpha = .01$

$z_{.01} = 2.326$

(4) $$z = \frac{(103-100)}{12/\sqrt{40}} = \frac{3}{1.8974} = 1.5811$$

(5) Reject H_0 if $|z| \geq z_\alpha$

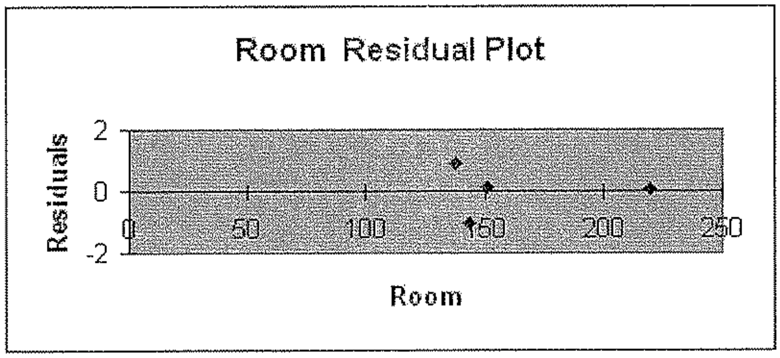
$1.5811 \geq 2.326$ FALSE

(6) Fail to Reject H_0; The data appear to support the location goal of more than 100 people passing by per hour based on a sample of 40 and a 1% level of significance.

9. Using the regression output below and a 10% level of significance, what recommendations and conclusions can be drawn?

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9416
R Square	0.8866
Adjusted R Square	0.8299
Standard Error	1.2317
Observations	4



ANOVA

	df	SS	MS	F	Significance F
Regression	1	23.71569	23.71569	15.63167	0.0584
Residual	2	3.034313	1.517157		
Total	3	26.75			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	40.7358	3.1005	13.1384	0.0057	27.3953	54.0762	27.3953	54.076
Room	0.0736	0.0186	3.9537	0.0584	-0.0065	0.1537	-0.0065	0.153

ANSWER: Based on the sample data the regression model is $\hat{y} = 40.7358 + 0.0736x$. Based on an examination of the p-value for the slope of the equation, the model appears to be significant at the 10% level of significance. The model also represents a good fit as evidenced by the goodness of fit measure of 0.8866. The goodness of fit measure (coefficient of determination) shows that 88.66% of the variability of the dependent variable can be explained by the independent variable. The regression model also shows that there is a strong positive relationship between the dependent variable and the independent variable. This is evidenced by the strength of relationship measure (correlation coefficient) of 0.9416. Finally the four assumptions that were made about the error term appear to be valid based on an examination of the residual plot.

10. Using an independent variable value of 200, what would the predicted value of the dependent variable be using the regression model from the table listed above?

$$\hat{y} = 40.7358 + 0.0736x$$

$$\hat{y} = 40.7358 + 0.0736 * 200$$

$$\hat{y} = 40.7358 + 14.72$$

$$\hat{y} = 55.4558$$

Using an independent variable value of 200, the predicted dependant variable would be 55,4558.