Practice Final Exam **STA 2023** **Name __________________________**
Section__________________________

1. **Identify the type of sampling used in each case.**
   (Random, Systematic, Stratified, Cluster, or Convenience)

   a. **A pollster selects drivers who are having their cars repaired at a local Sears auto store.** — a. Convenience

   b. **A pollster selects every 50th name in a telephone book.** — b. Systematic

   c. **A pollster selects 100 men and 100 women.** — c. Stratified

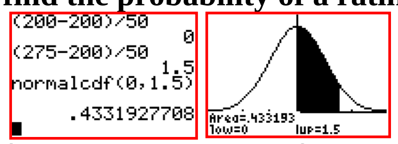
   d. **A pollster selects 50 people from each of 50 countries.** — d. Cluster

   e. **A pollster writes the names of each voter on a card, shuffles the cards, and then draws 25 names.** — e. Random

2. **A bank's loan officer rates applicants for credit. The ratings are normally distributed with a mean of 200 and a standard deviation of 50.**

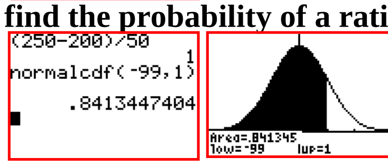
   a. **If an applicant is randomly selected, find the probability of a rating that is between 200 and 275.** — a. 0.4332

$$z = \frac{x-\mu}{\sigma} = \frac{275-200}{50} = +1.50$$



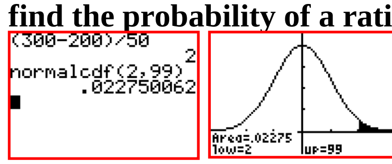
   b. **If an applicant is randomly selected, find the probability of a rating that is below 250.** — b. 0.8413

$$z = \frac{x-\mu}{\sigma} = \frac{250-200}{50} = +1.00$$



   c. **If an applicant is randomly selected, find the probability of a rating above 300.** — c. 0.0228

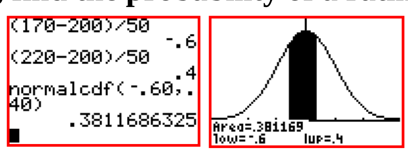
$$z = \frac{x-\mu}{\sigma} = \frac{300-200}{50} = +2.00$$



   d. **If an applicant is randomly selected, find the probability of a rating between 170 and 220.** — d. 0.3812

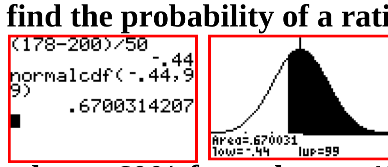
$$z = \frac{x-\mu}{\sigma} = \frac{170-200}{50} = -0.60$$

$$z = \frac{x-\mu}{\sigma} = \frac{220-200}{50} = +0.40$$



   e. **If an applicant is randomly selected, find the probability of a rating above 178.** — e. 0.6700

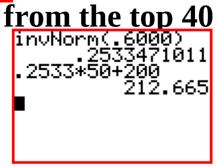
$$z = \frac{x-\mu}{\sigma} = \frac{178-200}{50} = -0.44$$



   f. **Find D_6, the score which separates the lower 60% from the top 40%.** — f. 212.67

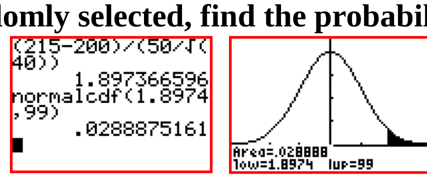
$$z = \frac{x-\mu}{\sigma} = \frac{x-200}{50} = .2533$$

$$x = .2533(50) + 200 = 212.67$$



   f. **If 40 different applicants are randomly selected, find the probability that their mean is above 215.** — g. 0.0287

$$z = \frac{x-\mu}{\sigma} = \frac{215-200}{50/\sqrt{40}} = 1.897$$



3. **A fund raising committee is to be selected from a group of 50 members, including 20 college graduates (12 of whom are women) and 30 people who did not graduate from college (14 of whom are women).**

a. **If the chairperson is randomly selected, find the probability of getting a woman, given that they have graduated from college.** P[W | G]= $\frac{P[W \cap G]}{P[G]} = \frac{12/50}{20/50} = 3/5 =$ a. __0.6000_

b. **If the chairperson is randomly selected, find the probability of getting a man or a college graduate.** P[M or G] = P[M] + P[G] – P[M∩G] $= \frac{24}{50} + \frac{20}{50} - \frac{8}{50} = \frac{36}{50} =$ b. __0.7200_

c. **If two different members are randomly selected for a special project, find the probability that they are both women.** P[both W] = $\frac{26}{50} \bullet \frac{25}{49} = \frac{13}{49} =$ c. __0.2653_

d. **At each meeting, one of the 50 members is randomly chosen to be secretary. Find the probability that the first two secretaries are both men.**

P[both M] = $\frac{24}{50} \cdot \frac{24}{50} = \frac{144}{625} =$ d. __0.2304_

e. **If the chairperson is randomly selected, find the probability that the chairperson is a male college graduate.** $P[M \cap G] = \frac{8}{50} =$ e. __0.1600_

f. **Are sex and college graduation mutually exclusive? (yes or no)** f. __ no ___

g. **Are sex and college graduation independent? (yes or no)** g. __ no ___

	Grad.	Non-Grad.	Totals
Males	8	16	24
Females	12	14	26
Totals	20	30	50

4. **An appliance manufacturing company obtained the following data on the length of time (in years) that 14 of their refrigerators operated before requiring repairs:**

4.0, 3.5, 5.8, 7.2, 7.8, 2.8, 0.8
6.1, 3.2, 2.9, 3.3, 1.6, 1.5, 2.4

Assuming the length of time is normally distributed, obtain a 98 percent confidence interval for the population mean.

$2.251 < \mu < 5.307$

98% C.I. for μ

$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$

$n = 14$

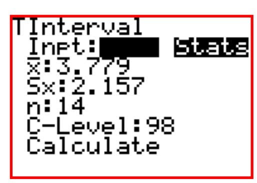
$\bar{x} = 3.779$

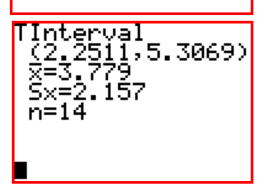
$s = 2.157$

$3.779 \pm 2.650 * \frac{2.157}{\sqrt{14}}$

3.779 ± 1.528

$\boxed{2.251 < \mu < 5.307}$





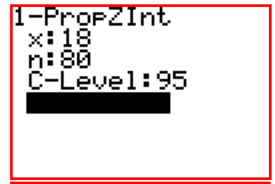
5. **A manufacturer of computer disk drives found that of the 80 drives selected at random from a very large production 18 were defective. Determine a 95 percent confidence interval for the population proportion of defective disk drives.**

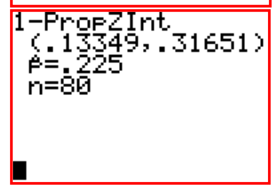
__0.133 < P < 0.317__

$$n = 80$$
$$x = 18$$
$$\hat{p} = \frac{x}{n} = \frac{18}{80} = 0.225$$
$$\hat{q} = 1 - \hat{p} = 1 - 0.225 = 0.775$$

__95% C.I. for p__

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}$$
$$0.225 \pm 1.96\sqrt{\frac{0.225 * 0.775}{80}}$$
$$0.225 \pm .092$$
$$\boxed{0.133 < p < 0.317}$$



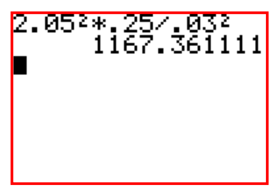


6. **A survey is to be conducted to estimate the proportion of U.S. citizens who feel that tariff restrictions should be imposed on foreign imports in the U.S. Determine how large the sample should be so that, with 96 percent confidence, the sample proportion will not differ from the true proportion by more than 0.03.**

__1168__

$$n = \frac{Z_{\alpha/2}^2 pq}{E^2} = \frac{(2.054)^2(0.25)}{(0.03)^2} = 1171.92$$

$\boxed{\text{Rounded up to 1172 citizens}}$



7. **A marketing survey involves product recognition in New York and California. Of 558 New Yorkers surveyed, 193 knew the product while 196 out of 614 Californians knew the product. At the 0.05 significance level, test the claim that the recognition rates are the same in both states.**

New York	California
$n_1 = 558$	$n_2 = 614$
$x_1 = 193$	$x_2 = 196$
$\hat{p}_1 = \frac{x_1}{n_1} = \frac{193}{558} = 0.346$	$\hat{p}_2 = \frac{x_2}{n_2} = \frac{196}{614} = 0.319$
$\hat{q}_1 = 1 - \hat{p}_1 = 0.654$	$\hat{q}_2 = 1 - \hat{p}_2 = 0.681$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{193 + 196}{558 + 614} = 0.332$$

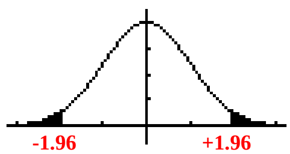
$$\bar{q} = 1 - \bar{p} = 1 - 0.332 = 0.668$$

(Claim) $p_1 = p_2 \longrightarrow$ $H_o: p_1 - p_2 = 0$

$H_1: p_1 - p_2 \neq 0$

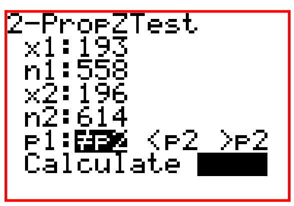
Test Statistic: $$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{(0.346 - 0.319) - 0}{\sqrt{\frac{(0.332)(0.668)}{558} + \frac{(0.332)(0.668)}{614}}} = 0.980$$

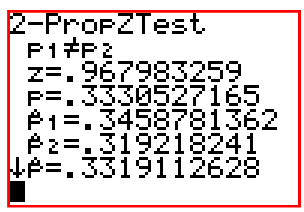
Test:

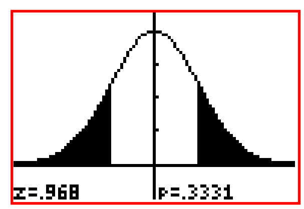


Fail to Reject H_o

Conclusion – Recognition rates are the same in both states.







8. **Construct a 99% confidence interval for the difference between the two populations proportions referred to in problem 8.**

__-0.044 < P < 0.098__

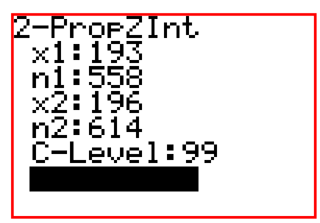
__99% C.I. for p_1-p_2__

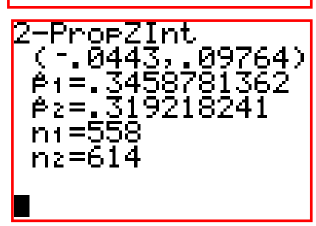
$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} * \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$(0.346 - 0.319) \pm 2.575 * \sqrt{\frac{0.346 * 0.654}{558} + \frac{0.319 * 0.681}{614}}$$

$$0.027 \pm 0.071$$

$$\boxed{-0.044 < p_1 - p_2 < 0.098}$$





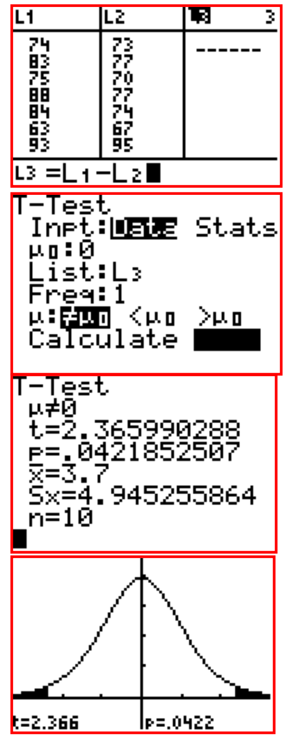
9. A test of abstract reasoning is given to a random sample of students before and after they completed a formal logic course. The results are given below. At the 0.05 significance level, test the claim that the mean score is not affected by the course.

Before	74	83	75	88	84	63	93	84	91	77
After	73	77	70	77	74	67	95	83	84	75
d =($x_1 - x_2$) =	1	6	5	11	10	-4	-2	1	7	2

$n = 10$

$\bar{d} = 3.7$

$s_d = 4.95$

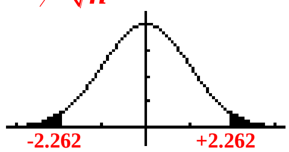


(Claim) H$_o$: $\mu_d = 0$

H$_1$: $\mu_d \neq 0$

Test Stat: $t = \dfrac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \dfrac{3.7 - 0}{4.95 / \sqrt{10}} = 2.366$

Test:



Reject H$_o$

Conclusion – The mean score is affected by the course.

10. A pollster interviews voters and claims that her process is a simple random selection. Listed below is the sequence of voters identified according to their sex. At a .05 level of significance, test her claim that the sequence is random according the criterion of sex. (Note: use the Sign Test).

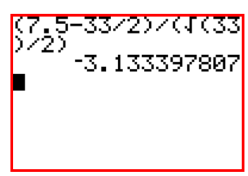
M M M M M F F F M M M
M M M F M M M M M M F
M F M M M M M F M M M

n = 33
x = 7

+	+	+	+	+	-	-	-	+	+	+
+	+	+	-	+	+	+	+	+	+	-
+	-	+	+	+	+	+	-	+	+	+

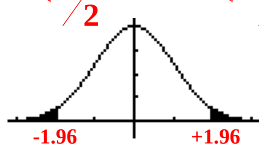
H$_o$: % of males = % of females

H$_1$: % of males ≠ % of females



Test Statistic: $Z = \dfrac{(x + 0.5) - n/2}{\sqrt{n}/2} = \dfrac{7.5 - 33/2}{\sqrt{33}/2} = -3.13$

Test:



Reject Ho

Conclusion – The sequence is not random by sex.

11. **In a study of crop yields, two different fertilizer treatments are tested on parcels with the same size and soil conditions in 10 different southern states. Listed below are the per acre yields for the sample plots.**
At the 0.05 significance level, and using the Wilcox Rank-Sum test, determine if there is a difference between the two treatments.

State	Fl.	Ga.	Ms.	La.	Ak.	NC	SC	Va.	WV	Tn.
Rank	2	4	7	16	5	8	9	6	1	3
Treatment A	132	137	142	160	139	143	145	140	131	136
Rank	10	11.5	14.5	14.5	11.5	19	17	18	20	13
Treatment B	148	152	159	159	152	167	163	165	180	156

$n_1 = 10$

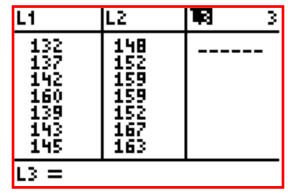
$n_2 = 10$

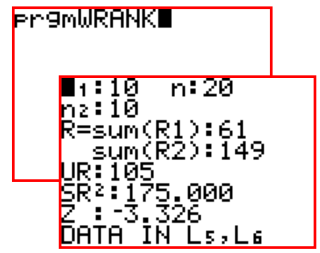
$\Sigma R_A = 61$

$\Sigma R_B = 149$

$$\mu_R = \frac{n_1(n_1+n_2+1)}{2} = \frac{10(10+10+1)}{2} = 105$$

$$\sigma_R = \sqrt{\frac{n_1 n_2(n_1+n_2+1)}{12}} = \sqrt{\frac{10*10(10+10+1)}{12}} = 13.23$$

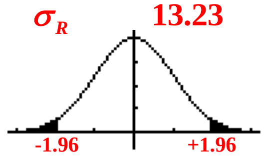




H_0: Treatments are the same
H_1: Treatments are different

Test Statistic: $Z = \dfrac{R-\mu_R}{\sigma_R} = \dfrac{61-105}{13.23} = -3.326$

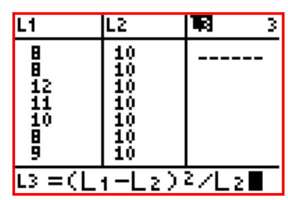
Test:

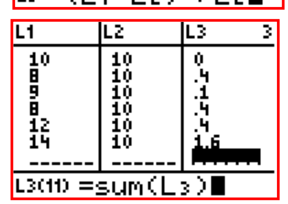


Reject Ho
Conclusion – <u>Treatments are different.</u>

12. **In the Decimal Representation of Π, the first 100 digits occur with the below listed frequencies. At a 0.05 significance level, test the claim that the digits uniformly distributed. (Note: Use the Chi-Square Goodness of Fit Test)**

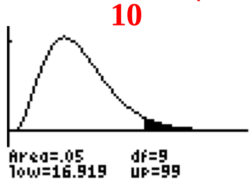
Digit	0	1	2	3	4	5	6	7	8	9	
Frequency	8	8	12	11	10	8	9	8	12	14	(Observed)
	10	10	10	10	10	10	10	10	10	10	(Expected)





H_0: Uniform distribution
H_1: Non-Uniform distribution

Test Statistic: $X^2 = \sum\dfrac{(O-E)^2}{E} = \dfrac{(8-10)^2}{10} + \dfrac{(8-10)^2}{10} + \dfrac{(12-10)^2}{10} + ... + \dfrac{(14-10)^2}{10} = 4.2$



Test:
df = k-1
Fail to Reject H_o

Conclusion – <u>The digits are uniformly distributed.</u>

13. **Two separate tests are designed to measure a student's ability to solve problems. Several students are randomly selected to take both test and the results are given below**

Test A	64	48	51	59	60	43	41	42	35	50	45
Test B	91	68	80	92	91	67	65	67	56	78	71

(a) **Plot the scatter diagram on the back of this page.**

(b) **Find the value of the linear correlation coefficient r.** **b.** .9745

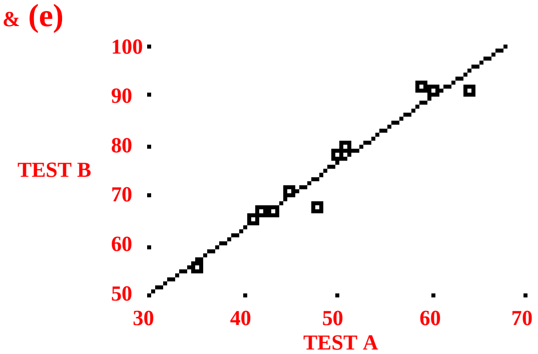
(c) **Test the significance of the correlation coefficient r at a 5% significance level**

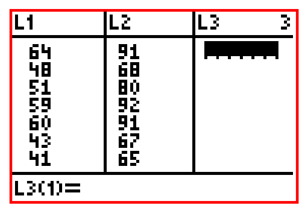
(d) **Use the given sample data to find the estimated equation of the regression line.** **d.** y = 1.3188x + 10.588

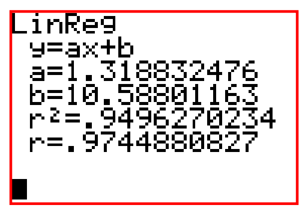
(e) **Plot the regression line with the equation given in part (d). Plot that line on the same scatter diagram given in part (a).**

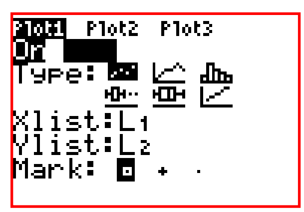
(f) **Predict a student's score on the Test B if he scored a 52 on Test A.** $y = 1.3188(52) + 10.5880 = 79.1656$ **f.** 79.1656 or 79

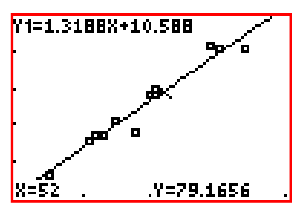
(a) & (e)









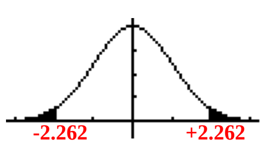


(c) $H_0: \rho = 0$ (there is no Linear Correlation)

$H_1: \rho \neq 0$ (there is a Linear Correlation)

Test Statistic: $t = \dfrac{r}{\sqrt{\dfrac{1-r^2}{n-2}}} = \dfrac{0.9745}{\sqrt{\dfrac{1-0.9745^2}{11-2}}} = 13.03$

Test:



Reject H_o

Conclusion – There is a Linear Correlation.