STATISTICS 641 - ASSIGNMENT 7

DUE DATE: NOON, MONDAY, NOVEMBER 22, 2021

Name _____

Email Address _____

Make this cover sheet the first page of your Solutions.

ASSIGNMENT 7 - DUE: NOON (CST), MONDAY, NOVEMBER 22, 2021

- Read Handout 12
- Supplemental Reading from Devore book: Chapters 8

P1. (10 points) A researcher presents you with the following data Y_i, $i = 1, 2, \ldots 29$, which is independent and identically normally distributed with a mean μ and variance σ^2.

1. Set up the rejection region for testing the hypotheses $H_0 : \mu = 20$ vs $H_1 : \mu \neq 20$, based on $Y_1, \ldots, Y_{29}$ and $\alpha = 0.05$. Assume σ is unknown.
2. Calculate the power of your test for the following values of the true parameter μ:

   19.9, 19.95, 19.99, 20, 20.05, 20.1, 20.15, 20.2, 20.25, 20.3, 20.4, 20.5.

   The researcher provides the estimate, $\hat{\sigma} = .43$. This value of σ should just be used in the power calculation and not in the actual testing procedure. Use your results to sketch a power curve for your test. Be sure to label your axes clearly.
3. Determine the necessary sample size so that the power at $\mu = 20.15$ is at least 0.80 for a level $\alpha = .05$ test of $H_0 : \mu \leq 20$ vs $H_1 : \mu > 20$.

P2. (10 points) A new additive has been formulated to reduce the reaction time in a chemical process. With the previously used additive, the average reaction time was 10 minutes. In order to evaluate the effectiveness of the new additive, 15 batches of the material are formulated and the new additive is placed in the batches. From previous studies, reaction times appear to have a normal distribution.

1. The mean reaction time from the 15 batches was 8.7 minutes with a standard deviation of 2 minutes. Is there significant evidence using an $\alpha = .01$ test that the average reaction time has been reduced? Include the p-value with your decision.
2. The process engineer had claimed that the new additive will reduce the average reaction time by at least 1.5 minutes. What is the probability that the experiment will be able to detect a reduction of the average reaction to 8.5 minutes or smaller using $\alpha = .01$?
3. A new study is to be designed. What sample size is needed for an $\alpha = .05$ test to have at least an 80% chance to detect that the average reaction time is 9 minutes or less?

P3. (10 points) A new device has been developed which allows patients to evaluate their blood sugar levels. The most widely device currently on the market yields widely variable results. The new device is evaluated by 25 patients having nearly the same distribution of blood sugar levels yielding the following data:

125	123	117	123	115	112	128	118	124
111	116	109	125	120	113	123	112	118
121	118	122	115	105	118	131		

1. Is there significant evidence ($\alpha = .10$) that the standard deviation in the readings from the new device is less than 10?
2. Compute the probability of a Type II error in using your test from part 1. for the following values of σ: 5, 6, 7, 8, 9, 10
3. Construct an upper 90% confidence bound on the standard deviation of the new device. Is this bound consistent with your answer to the question in part 1.?

P4. (10 points) Refer to the blood sugar device data in Problem 3.

1. Is there significant ($\alpha = .05$) evidence that median blood sugar readings was less than 120 in the population from which the 25 patients were selected? Use the sign test and report the p-value.
2. Is there significant ($\alpha = .05$) evidence that median blood sugar readings was less than 120 in the population from which the 25 patients were selected? Use the Wilcoxon signed rank test and report the p-value.
3. Place a 90% upper bound on the median blood sugar reading.

P5. (10 points) The current method of identifying patients at risk of sudden cardiac death can be identified with 80% accuracy. A change in the method has hopefully improved the accuracy. To evaluate the new method, 100 people are tested and the new method produced the result on 92 of the 100 people.

1. Place a 95% confidence interval on the accuracy of the device.
2. Is there substantial evidence ($\alpha = .05$) that the improved method has increased the accuracy over the current method?
3. Compute the power of the test in part 2. to detect that the accuracy of the improved method is 75%, 80%, 85%, 90%, 95%.
4. How many patients would need to be included in a new study in order to have a power value of 80% if the new method had an accuracy of 90%.

P6. **Multiple Choice (50 points) SELECT ONE** of the following letters **(A, B, C, D, or E)** corresponding to the **BEST** answer. Show details for partial credit.

(MC1.) Suppose that $X_1, \cdots, X_n$ are to be used to construct a 95% prediction interval for a normal population. The researcher notes that the data was collected by an automatic sampler which may result in $X_1, \cdots, X_n$ having a high positive correlation. If the prediction interval was computed using the formula: $\bar{X} \pm t_{.025,n-1} S\sqrt{1+\frac{1}{n}}$, the resulting interval

A. will have a level of confidence less than 95%.
B. will have a level of confidence greater than 95%.
C. will have a level of confidence equal to 95%.
D. will be too wide.
E. none of the above

(MC2.) In a level $\alpha = .05$ test of $H_o : \mu \leq 17$ versus $H_1 : \mu > 17$, where μ is the mean of a normally distributed population, the sample size needed to have a Type II error rate of at most 0.10 whenever $\mu > 17 + .5 * \sigma$ is

A. 13
B. 22
C. 35
D. 70
E. need the non-central t cdf in order to determine sample size

(MC3.) The power of a test of the hypothesis: $H_1 : \mu < \mu_o$

A. is one minus the probability of a Type II error at μ_o
B. is the probability of a Type II error
C. varies depending on the value of μ

D. is not a function of the value of α

E. none of the above

(MC4.) In testing the hypotheses $H_o : \sigma \leq 23.8$ versus $H_1 : \sigma > 23.8$, where σ is the standard deviation of a normally distributed population, an $\alpha = .05$ test was run using a independent random sample of size $n = 10$. The probability of a Type II error when $\sigma = 47.9$ is

A. .05

B. .10

C. .90

D. .95

E. need noncentral Chi-squared tables to compute power

(MC5.) A researcher wants to determine if there is an increase in the likelihood that people will purchase a product after a redesign of the product. The current market share is 20%. Initially, the researcher was planning on using a random sample of n=20 persons with an $\alpha = .05$ test to evaluate the product. He wants you to calculate the chance that the study will fail to detect that preference for the product has been increased if in fact the preference for the new product is 40%. This chance is

A. .316

B. .416

C. .596

D. .950

E. cannot be determined with the given information

(MC6.) A random sample of n=15 from a normally distributed population is used to construct a level $\alpha = .01$ test of $H_o : \mu \leq 20$ versus $H_1 : \mu > 20$, where μ is the mean of the population. The probability of a Type II error for $\mu > 20 + .8\sigma$ is at most

A. .05

B. .22

C. .32

D. .55

E. cannot be determined from the given information

(MC7.) A psychologist is investigating the IQ level of young children who have been in a head start program. She wants to determine if the variation in IQ scores for the population of head start students is smaller than the variation in the general population of children under the age of 6 which has a variation of $\sigma = 10.2$. She also informs you that the distribution of IQ scores is highly right skewed. Suppose she uses the test: reject H_o is $\frac{(n-1)S^2}{(10.2)^2} < \chi^2_{.95,n-1}$, where S is the standard deviation from a random sample of n head start students, to test whether σ is less than 10.2 with an α value of 0.05.

A. the actual level of significance will be greater than 0.05.

B. the actual level of significance will be less than 0.05.

C. the actual level of significance will be very close to 0.05.

D. the actual level of significance will be exactly 0.05.

E. it is impossible to determine the effect of skewness on the actual level of significance.

Use the following information for Problems MC8-MC9.

A process engineer samples a continuous flow of the company's product 200 times per day and obtains the following pH levels in the product : $X_1, \cdots, X_{200}$. He determines that the daily pH levels are related by $X_t = \theta + \rho X_{t-1} + e_t$, where the e_ts have independent $N(0, \sigma_e^2)$ distributions and $\rho \approx .92$.

(MC8.) The engineer constructs a nominal 95% confidence interval for the average daily pH level, μ, using the formula $\bar{X} \pm t_{.025,199}(s/\sqrt{200})$, where $\bar{X}$ and s are the sample mean and standard deviation for a given days pH levels. The true coverage probability of this confidence interval

A. is 0.95.

B. is much less than 0.95.

C. is very close to 0.95.

D. is much greater than 0.95.

E. may be greater than 0.95 or less than 0.95 depending on the distribution of the X_t's.

(MC9.) Refer to Problem MC8. The nominal pH level of the product is 5.3. The process engineer wants to test if the pH on a given day is different from 5.3. He uses $t = \frac{\bar{X}-5.3}{s/\sqrt{200}}$ as his test statistic. Next, he uses the t-distribution with df=199 to compute the p-value of the observed data. The computed p-value will be

A. correct because the sample size is large.

B. much smaller than the correct p-value.

C. much larger than the correct p-value.

D. very close to the correct p-value because the sample size is large.

E. may be greater or less than the correct value depending on the size of σ.

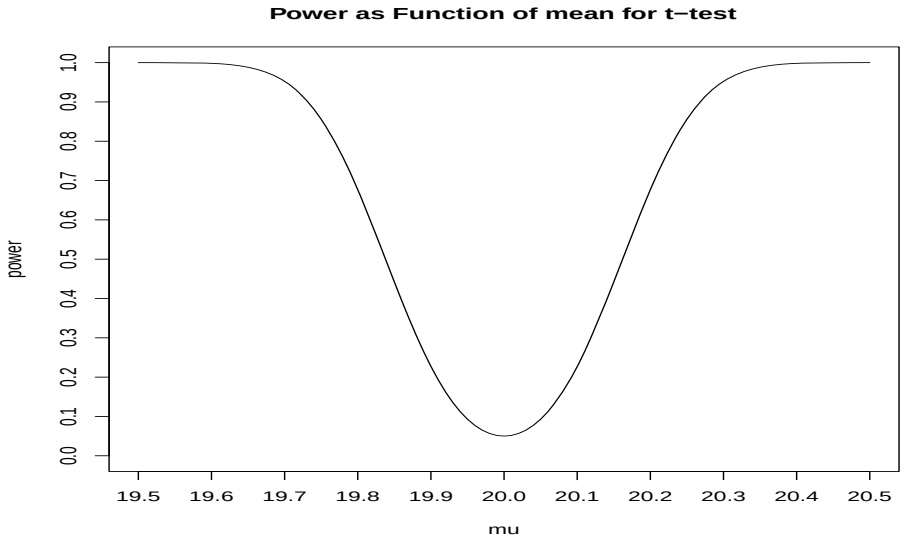
STAT 641 - ASSIGNMENT 7 - SOLUTIONS - Fall 2021

P1. (10 points)

Let $Y_1, Y_2, \cdots, Y_{29} \sim iid\ N(\mu, \sigma^2)$.

1. $H_0 : \mu = 20$ vs $H_1 : \mu \neq 20$, $\alpha = 0.05$.
   i. Test statistic is
   $$t = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}} = \frac{\bar{Y} - 20}{s/\sqrt{29}} \sim t_{28}.$$
   ii. Decision rule: reject H_0 if $|t| \geq t_{0.025,28} = 2.048$.
2. The power function is given by $\gamma(\mu) = P\left(\left|\frac{\bar{Y}-\mu_0}{s/\sqrt{n}}\right| \geq t_{0.025,28} \middle| \ \mu\right) =$
   $$1 - F(t_{0.025,28}) + F(-t_{0.025,28}) = 1 - pt(qt(0.975, 28), 28, ncp) + pt(qt(0.025, 28), 28, ncp)$$
   where F is a cdf of noncentral t distribution with $df = 28$ and noncentrality parameter
   $$ncp = \frac{\sqrt{n}(\mu - \mu_0)}{\hat{\sigma}} = \frac{\sqrt{29}(\mu - 20)}{.43}$$
   :

```
      mu       ncp     power
 [1,] 19.50 -6.261820 0.99997712
 [2,] 19.60 -5.009456 0.99795710
 [3,] 19.70 -3.757092 0.95208439
 [4,] 19.75 -3.130910 0.85586759
 [5,] 19.80 -2.504728 0.67658693
 [6,] 19.85 -1.878546 0.44207281
 [7,] 19.90 -1.252364 0.22729806
 [8,] 19.95 -0.626182 0.09286722
 [9,] 20.00  0.000000 0.05000000
[10,] 20.05  0.626182 0.09286722
[11,] 20.10  1.252364 0.22729806
[12,] 20.15  1.878546 0.44207281
[13,] 20.20  2.504728 0.67658693
[14,] 20.25  3.130910 0.85586759
[15,] 20.30  3.757092 0.95208439
[16,] 20.40  5.009456 0.99795710
[17,] 20.50  6.261820 0.99997712
```



3. For given $\alpha = 0.05$ and $H_1 : \mu > 20$, we want to find n such that $\gamma(20.15) \geq 0.8$. Then,

Approach 1:

$$\begin{aligned}\gamma(20.15) &= P\left(\frac{\bar{Y}-\mu_0}{s/\sqrt{n}} \geq t_{0.05,n-1}\middle|\ \mu = 20.15\right) \\ &= 1 - P\left(t(n-1,d) \leq t_{0.05,n-1}\right) \\ &= 1 - F(t_{0.025,n-1})\end{aligned} \tag{1}$$

where F is the cdf of noncentral t distribution with df-1 and noncentrality $d = \frac{\sqrt{n}(\mu_1-\mu_0)}{\widehat{\sigma}} = \frac{\sqrt{n}(20.25-20)}{.43}$. Now, find n iteratively such that $1 - F(t_{0.05,n-1}) \geq 0.8$. Using the following R program we obtain:

```
s=.43
a = .05
m0 = 20
m1 = 20.15
n = seq(50,55,1)
df = n-1
d = sqrt(n)*(m1-m0)/s
p = 1-pt(qt(1-a,n-1),n-1,d)
data = cbind(n,df,d,p)
data
```

```
      n df        d         p
[1,] 50 49 2.466652 0.7845072
[2,] 51 50 2.491196 0.7917195
[3,] 52 51 2.515501 0.7987172
[4,] 53 52 2.539573 0.8055052
[5,] 54 53 2.563419 0.8120886
[6,] 55 54 2.587046 0.8184721
```

Thus, the necessary sample size is greater than or equal to 53.

Approach 2: Using Table A11 on page 29 in Handout 12 with $\phi = \frac{|\delta|}{\sigma} = \frac{.15}{.43} = .35$, $\beta = .2$ $\alpha = .05$ Single-Sided Test, we obtain from the table $n = 52$ confirming the value obtained in Approach 1. This value would result in a test having power 0.7987 using the R function:

```
power.t.test(n=52,delta=.15,sd=.43,sig.level=.05,power=,type="paired",alternative=c("one.sided"))
```

Approach 3: Assume σ is known and use the formula in Handout 12 knowing that the obtained value would be somewhat of an underestimate:

$$n = \frac{\sigma^2(z_{.05} + z_{.2})^2}{(\mu - \mu_o)^2} = \frac{(.43)^2(1.645 + 0.8416)^2}{(20 - 20.15)^2} = 50.8$$

Thus, the necessary sample size is greater than or equal to 51.
This value would result in a test having power 0.792 using the R function:

```
power.t.test(n=51,delta=.15,sd=.43,sig.level=.05,power=,type="paired",alternative=c("one.sided"))
```

Approach 4: Using the above function with n blank and power=.8, we obtain n=52.2 which implies n=53:

```
power.t.test(n=,delta=.15,sd=.43,sig.level=.05,power=.8,type="paired",alternative=c("one.sided"))
```

Based on all the above computations, n = 53 would be the required sample size although n = 52 yields a power of .7987 which is very close to .80.

P2. (10 points) Let μ be the reaction time in a chemical process using the new additive.

Test the hypotheses: $H_o : \mu \geq 10$ vs $H_1 : \mu < 10$

1. From the n=15 batches: $\bar{Y} = 8.7$ and $S = 2$. Using $\alpha = .01$,

Reject H_o if $\bar{Y} < 10 - t_{.01}S/\sqrt{n} = 10 - (2.624)(2)/\sqrt{15} = 8.645 \Rightarrow \bar{Y} = 8.7 > 8.455$ and

$p - value = P[t_{14} < \sqrt{15}(8.7 - 10)/2] = pt(-2.517, 14) = .0123 > .01 = \alpha \Rightarrow$

Fail to reject H_o and conclude there is not sufficient evidence that the average reaction time has been reduced using the new additive.

2. Using $\sigma \approx 2$, compute the power at $\mu = 8.5$:

   $\gamma(8.5) = \text{P[reject } H_o \text{ at } \mu = 8.5] = P[t < -t_{.01,14}] = P[t < -2.624]$

   where t has a non-central t-distribution with df = 14 and non-centrality parameter,

   $\Delta = \sqrt{15}(8.5 - 10)/2 = -2.9047$. Therefore,

   the power of the test at $\mu \leq 8.5$ is given by $\gamma(8.5) \geq pt(-2.624, 14, -2.9047) = .6165$

   Using the R-fucntion, we obtain power = .6163

   ```
   power.t.test(n=15,delta=1.5,sd=2,sig.level=.01,power=,type="paired",alternative=c("one.sided"))
   ```

3. Using the table on page 29 of Handout 12 with a one-sided test having $\alpha = .05$; $\beta = 1 - .80 = .2$; $\phi \approx |9 - 10|/2 = .5$; we have n=27.

   The actual power using n=27 observations is computed using

   $-t_{.05,26} = -1.7056$, and $\Delta = \sqrt{27}(9 - 10)/2 = -2.5981$ which yields

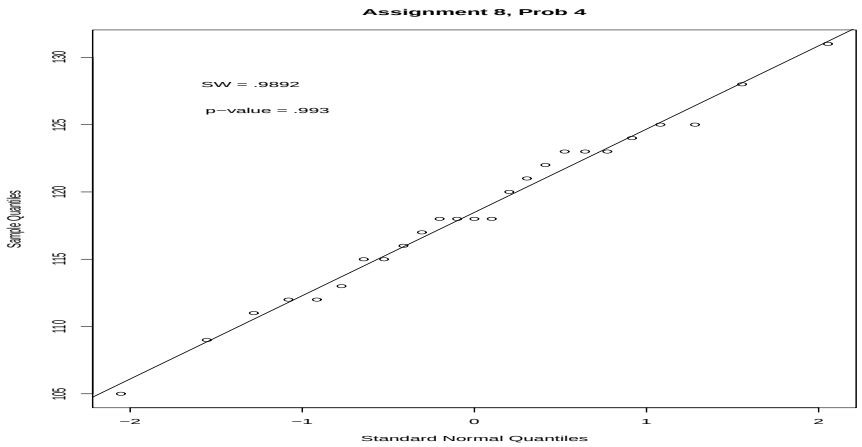
   power at $\mu = 9$ is $\gamma(9) = pt(-1.7056, 26, -2.5981) = .81$

   The R-function can also be used

   ```
   power.t.test(n=,delta=1,sd=2,sig.level=.05,power=.8,type="paired",alternative=c("one.sided")) yields
   n=26.1 round to n=27
   power.t.test(n=27,delta=1,sd=2,sig.level=.05,power=,type="paired",alternative=c("one.sided")) yields
   power = 0.81
   ```

P3. (10 points) The following normal reference distribution plot along with a p-value of .993 from the Shapiro-Wilk test indicates that a normal distribution provides an excellent fit to the data with

$n = 25, \quad \bar{X} = 118.48, \quad S = 6.1922.$



1. Test the hypotheses: $H_o : \sigma \geq 10$ versus $H_1 : \sigma < 10$ with rejection region:

   Reject H_o if $(n-1)S^2/(10)^2 \leq \chi^2_{1-.10,24} = \chi^2_{.90,24} = 15.659$, where $\chi^2_{.90,24}$ is the .9 upper percentile.

   From the data, $(n-1)S^2/(10)^2 = (25-1)(6.192)^2/(10)^2 = 9.20 < 15.659 \Rightarrow Reject H_o$ and conclude there is significant evidence that the new device produces readings which have a standard deviation less than 10.

p-value $= P[\chi^2_{24} \leq 9.20] = pchisq(9.20, 24) = .003$ which is less than $\alpha = 0.10$

2. $\beta(\sigma_1) = P[\text{Type II error at } \sigma_1] = P[\chi^2_{24} \geq \frac{(10)^2}{\sigma_1^2} 15.659] = 1 - pchisq\left(\frac{(10)^2}{\sigma_1^2}\ 15.659\right)$

σ_1	5	6	7	8	9	10
$\beta(\sigma_1)$	.0000269	.00872	.128	.435	.734	0

$\beta(10) = 0$ because 10 is in the null space and hence a Type II error cannot occur at $\sigma = 10$.

3. From $P\left[\frac{(n-1)S^2}{\sigma^2} \geq \chi^2_{n-1,1-.1}\right] = P\left[\frac{(n-1)S^2}{\sigma^2} \geq 15.659\right] = .9$, where $\chi^2_{n-1,.9} = qchisq(1 - .9, 24) = 15.659$, we have that an upper 90% confidence bound on the standard deviation of the new device is given by
$\frac{\sqrt{n-1}S}{\sqrt{\chi^2_{n-1,.9}}} = \frac{\sqrt{25-1}(6.1922)}{\sqrt{15.659}} = 7.666$. Thus, we are 90% confident that σ is less than 7.666 which would be consistent with our conclusion that the data indicated that σ was less than 10.

P4. (10 points) Let $\tilde{\mu}$ be the median reading for the distribution of blood sugar device readings.

Test $H_o : \tilde{\mu} \geq 120$ versus $H_1 : \tilde{\mu} < 120$

1. **Sign Test**: One of the data values is 120, therefore delete it and use a sample size of $n^* = 25 - 1 = 24$.

   Let S_+ be the number of readings in the data greater than 120: The decision rule is

   Reject H_o if $S_+ \leq 7$, because $P[B \leq 7] = pbinom(7, 24, .5) = .032 < .05$ and $P[B \leq 8] = .076 > .05$, where B has a Binomial(n=24,p=.5) distribution.

   From the data, $S_+ = 10 > 7$ therefore, conclude there is not significant evidence that the median is less than 120.

   p-value $= P[B \leq S_+] = P[B \leq 10] = pbinom(10, 24, .5) = 0.271 > 0.05 = \alpha$

2. **Wilcoxon signed rank test**: Let W_+ be the sum of the ranks associated with the positive values of $X_i = Y_i - 120$:

   One of the values of X was 0, so we delete that observation and use $n^* = 25 - 1 = 24$.

   Reject $H_o : \tilde{\mu} \geq 120$ if $W_+ \leq qsignrank(.05, 24, TRUE) = 92$

   Table A.10 in Tables for STAT 641 is for $n \leq 20$. Using the asymptotic approximation,

   $W_{.05,24} \approx \frac{24(24+1)}{4} - Z_{.05}\sqrt{\frac{24(24+1)(48+1)}{24}} = 92.425$

   which is fairly close to the actual .05 percentile of 92 obtained from R.

   From the data, we have the sum of the ranks of $|X_i|$ associated with the positive values of X_i, that is, the values greater than 120, is $W_+ = 112.5 > 92$, therefore, fail to reject H_o and conclude there is not significant evidence that the median is less than 120.

   p-value $= P[W_+ \leq 112.5] = psignrank(112.5, 24, TRUE) = .151 > .05 = \alpha$

   Using the asymptotic approximation,

$$p - value \approx P\left[Z \leq \frac{112.5 - \frac{24(24+1)}{4}}{\sqrt{\frac{24(24+1)(48+1)}{24}}}\right] = P[Z \leq -1.071429) = 0.1420$$

   Using the following R function with x containing the 25 data values and c=rep(120,25), the R-function,

   ```
   wilcox.test(x,c,alternative="less",paired=TRUE)
   ```

   yields the following:

```
        Wilcoxon signed rank test with continuity correction

data:  x and c
V = 112.5, p-value = 0.1447
alternative hypothesis: true location shift is less than 0
```

- The sign test yielded p-value = .271; the Wilcoxon signed rank test yielded p-value = .145; and the t-test has a p-value of .116

  This is consistent with the result that when the population distribution is normally distributed, the t-test is the most powerful test.

3. There is strong evidence that the population distribution is a normal distribution, therefore a 90% upper bound on the median blood sugar reading is given by $\bar{Y}+t_{.10,24}S_Y/\sqrt{25} = 118.48+(1.318)(6.1922)/\sqrt{25} = 120.1$

   A distribution-free upper bound would be $(0,\ X_{(s)})$ where s is the smallest integer such that $.90 = pbinom(s-1, 25, .5)$ which yields s=17 with coverage=.946. Therefore, the distribution-free upper bound would be $(0,\ X_{(17)}) = (0,\ 122)$ which is somewhat larger than the normal based upper bound partially due to the higher coverage, .946 vs .90

   Note that the interval (0, 122) contains 120 and hence we would fail to reject H_o using the criterion that the feasible region for Q(.5) is the region (0, 122) which contains 120.

P5. (10 points) Let p be the probability of identifying patients at risk of sudden cardiac death using the new method. From the data, $\hat{p} = y/n = 92/100 = .92$

1. Because $min(n\hat{p}, n(1-\hat{p}) = 8 > 5$ and $n = 100 > 40$, the Agresti-Coull C.I. is appropriate. The Clopper-Pearson C.I. is given by $(C_L,\ C_U)$ where

$$C_L = \frac{1}{1+\frac{9}{92}F_{18,184,.025}} = 0.848; \qquad C_U = \frac{\frac{93}{8}F_{186,16,.025}}{1+\frac{93}{8}F_{186,16,.025}} = .965$$

   The 95% Agresti-Coull C.I. for p is given by $\tilde{p} \pm Z_{.025}\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}$ where
   $\tilde{n} = n + Z^2_{.025} = 100 + (1.96)^2 = 103.8416$, $\tilde{p} = (Y + .5Z^2_{.025})/\tilde{n} = (92 + .5(1.96)^2)/(103.8416) = .904$
   The 95% C.I. on p is $.904 \pm 1.96\sqrt{(.904)(1-.904)}/\sqrt{103.8416} = .904 \pm .0565 = (.848,\ .961)$
   With $min(n\hat{p}, n(1-\hat{p}) > 5$ and $n > 40$, the Agresti-Coull confidence interval is nearly identical to the Clopper-Pearson confidence interval.

2. Test the hypotheses $H_o : p \leq .8$ versus $H_1 : p > .8$ at the $\alpha = .05$ level.

   Let Y be the number of patients that were identified as being at risk out of the 100 patients.

   Reject H_o if $Y \geq B_{.05,100,.8} = qbinom(1-.05, 100, .8) = 86$.

   Need to check the size of the test:

   $\alpha = P[Y \geq 86] = 1 - P[Y \leq 85] = 1 - pbinom(72, 100, .8) = .0804 > .05$

   Change Rejection Region to Reject H_o if $Y \geq 87$ which yields

   $\alpha = P[Y \geq 87] = 1 - P[Y \leq 86] = 1 - pbinom(86, 100, .8) = .047 < .05$

   From the data, $Y = 92 > 87$. Thus, reject H_o and conclude that there is significant evidence that the new method has increased the accuracy relative to the old method.

   Let B have a Binomial(n=100,p=.8) distribution, then

   $p - value = P[B \geq 92] = 1 - pbinom(91, 100, .80) = .0009 < .05 = \alpha$

- Because $min[np_o, n(1-p_o)] = min[(100)(.8), (100)(.2)] = 20 > 5$, the asymptotic test could be used. It has p-value

$$p - value = P\left[Z \geq \frac{(\hat{p}-p_o)}{\sqrt{\frac{p_o(1-p_o)}{n}}}\right] = P\left[Z \geq \frac{(.92-.8)}{\sqrt{\frac{.8(1-.8)}{100}}}\right] = 1 - pnorm(3) = .0013$$

Note, that the value of the p-value from the asymptotic approximation is close to the exact value from the binomial distribution.

3. The power of the test in part (b.) is given by $\gamma(p) = P[\text{Reject } H_o] = P[Y \geq 87] = 1 - P[Y \leq 86]$

   $\gamma(p) = 1 - pbinom(86, 100, p)$ for p= .75, .80, .85, .90, .95

p	.75	.8	.85	.9	.95
$\gamma(p)$	0.002458	0.04691	0.3474	0.87611	0.9995

   The value of $\gamma(.8) = .04803$, demonstrates that the actual size of the test is .04803 not .05.

4. Using the code at the end of this document, I tried a range of sample sizes and computed exact sizes and powers using the Binomial distribution. In all cases, the actual size at the critical value obtained by `qbinom(0.95, n, 0.8)` was slightly greater than $\alpha = 0.05$, so I incremented the critical values by one. For example, with $n = 82$, `qbinom(0.95, 82, 0.8)` equals 71, but the size of the test with a critical value of 71 is `1 - pbinom(71 - 1, 82, 0.8)` $= 0.0836 > 0.05$. Incrementing the critical value by one gives `1 - pbinom(71 + 1 - 1, 82, 0.8)` $= 0.0458 < 0.05$. It turns out that $n = 82$ is the smallest n such that the size is less than 0.05 and the power is greater than 0.80.

   Alternatively, we could use the formula for the large-sample test. The required sample size n to achieve $\beta(.9) = 1 - \gamma(.9) = 1 - .8 = .2$ using an $\alpha = .05$ large-sample test is given by

$$n = \left[\frac{Z_\alpha\sqrt{p_o(1-p_o)} + Z_\beta\sqrt{p_1(1-p_1)}}{\delta}\right]^2 = \left[\frac{1.645\sqrt{.8(1-.8)} + .84\sqrt{.9(1-.9)}}{(.8-.9}\right]^2 = 82.8$$

   This suggests that n=83 is required to achieve the stated goals. We'll go with $n = 82$, since that was based on exact calculations.

P6. Multiple Choice (50 points)

(MC1.) **A.** The P.I. will be too narrow and hence will have a level of confidence less than 95%.

(MC2.) **C.** $n = \frac{\sigma^2(1.645+1.28)^2}{(.5\sigma)^2} = 34.2$

(MC3.) **C.** The power, $\gamma(\mu)$, is a function of μ

(MC4.) **B.** $\beta(47.9) = P\left[\chi^2_9 \leq \frac{(23.8)^2}{47.9)^2}16.919\right] = pchisq(4.177, 9) = .101$

(MC5.) **B.** Test the hypotheses $H_o : p \leq .2$ versus $H_1 : p > .2$ at the $\alpha = .05$ level.

Initially, Reject H_o if $Y \geq qbinom(.95, 20, .2) = 7$

Check level: $\alpha = P[Y \geq 7] = 1 - P[Y \leq 6] = 1 - pbinom(6, 20, .2) = .087 > .05$

Need to modify Rejection Region

Reject H_o if $Y \geq 8$ then $\alpha = P[Y \geq 8] = 1 - P[Y \leq 7] = 1 - pbinom(7, 20, .2) = .032 < .05$

$\beta(.4) = P[\text{Fail to Reject}H_o] = P[Y < 8] = P[B \leq 7] = pbinom(7, 20.4) = .416$

(MC6.) **C.** The test statistic would be $t = \sqrt{15}(\bar{Y} - 20)/S$ which has a non-central t-distribution with non-centrality parameter $\Delta = \frac{\sqrt{n}(\mu_1 - 20)}{\sigma} = \frac{\sqrt{15}(20+.8\sigma-20)}{\sigma} = .8\sqrt{15}$

$\beta(20 + .8\sigma) = P[t_{14,\Delta} < t_{.01,14}] = pt(2.6245, 14, .8\sqrt{15}) = .32$

Alternatively, if you use the a Z-test: Reject H_o if $\bar{Y} \geq 20 + 2.33\sigma/\sqrt{15}$

$\beta(20 + .8\sigma) = P[\bar{Y} < 20 + 2.33\sigma/\sqrt{15}$ when $\mu = 20 + .8\sigma] = P[Z < 2.33 + .8\sqrt{15}] = pnorm(-.768) = .22$

(MC7.) **A.** See the discussion on page 54 in Handout 12

(MC8.) **B.** See the discussion on page 35 in Handout 12

(MC9.) **B.** See the discussion on page 35 in Handout 12

```
####
#### (1)
####

##
## (2)
##

mu_0 <- 20
mu <- c(19.9, 19.95, 19.99, 20, 20.05, 20.1, 20.15, 20.2, 20.25, 20.3, 20.4, 20.5)
n <- 29
sigma_0 <- 0.43

Delta <- sqrt(n) * (mu - mu_0) / sigma_0
t_crit <- qt(0.975, n - 1)
gamma <- pt(-t_crit, n - 1, Delta) + 1 - pt(t_crit, n - 1, Delta)

plot(mu, gamma, xlab = expression(mu), ylab = "Power", type = "l", yaxt = "n")
abline(0.05, 0, lty = 2)
axis(2, at = c(0.05, 0.2, 0.4, 0.6, 0.8, 1.0))

##
## (3)
##

alpha <- 0.05
beta <- 0.2
phi <- abs(20.15 - mu_0) / 0.43

power.t.test(n = , delta = phi * 0.43, sd = 0.43, sig.level = alpha, power = 1 - beta,
  type = "one.sample", alternative = "one.sided")

## Linear interpolation of table values
w <- (phi - 0.30) / (0.35 - 0.30)
n_tbl <- 71 * (1 - w) + 52 * w

## Iterative solution using power function
n_seq <- 20:200; k <- length(n_seq)
gamma_seq <- numeric(k)
for(i in 1:k) {
  Delta_seq <- sqrt(n_seq[i]) * (20.15 - mu_0) / sigma_0
  t_crit_seq <- qt(0.95, n_seq[i] - 1)
  gamma_seq[i] <- 1 - pt(t_crit_seq, n_seq[i] - 1, Delta_seq)
}

####
#### (2)
####

mu_0 <- 10
n <- 15

##
```

```
## (1)
##

x_bar <- 8.7
s <- 2

## H_0: mu >= mu_0 vs H_a: mu < mu_0
t_stat <- sqrt(n) * (x_bar - mu_0) / s
qt(0.01, n - 1)
pt(t_stat, n - 1)

##
## (2)
##

Delta <- sqrt(n) * (8.5 - 10) / s
gamma <- pt(qt(0.01, n - 1), n - 1, Delta)

power.t.test(n = n, delta = 1.5, sd = s, sig.level = 0.01, power = ,
  type = "one.sample", alternative = "one.sided")

##
## (3)
##

alpha <- 0.05
beta <- 0.2
phi <- abs(9 - mu_0) / s

power.t.test(n = , delta = phi * s, sd = s, sig.level = 0.05, power = 0.8,
  type = "one.sample", alternative = "one.sided")

####
#### (3)
####

x <- c(125, 123, 117, 123, 115, 112, 128, 118, 124, 111, 116, 109, 125, 120, 113, 123,
  112, 118, 121, 118, 122, 115, 105, 118, 131)
n <- length(x)

x_bar <- mean(x)
s <- sd(x)

##
## Check whether data are Normally distributed
##

## Normal reference distribution plot
u <- (1:n - 0.5) / n
Q <- pnorm(u)
x_sort <- sort(x)

plot(Q, x_sort, xlab = "Normal Quantiles", ylab = "Sample Quantiles",
```

```
  main = "Normal Reference Distribution Plot")
abline(lm(x_sort ~ Q))

## Shapiro-Wilks test
shapiro.test(x)

##
## (1)
##

test_stat <- (n - 1) * s ^ 2 / 100
qchisq(0.10, n - 1)
p_value <- pchisq(test_stat, n - 1)

##
## (2)
##

s_alt <- 5:10
gamma <- pchisq((100 / s_alt ^ 2) * qchisq(0.1, n - 1), n - 1)
1 - gamma

##
## (3)
##

## Upper bound
sqrt((n - 1) * s ^ 2 / qchisq(0.10, n - 1))
sqrt((n - 1) * s ^ 2 / qchisq(0.90, n - 1))


####
#### (4)
####

##
## (1)
##

## Delete the instance of x = 120 and adjust sample size
x_star <- x[x != 120]
n_star <- length(x_star)

S_plus <- sum(x_star > 120)
p_value <- pbinom(S_plus, n_star, 0.5)

##
## (2)
##

y <- x_star - 120
y_ranked <- rank(abs(y))
W_plus <- sum(y_ranked[y > 0])
```

```
p_value <- psignrank(W_plus, n_star, TRUE)
wilcox.test(x, rep(120, n), alternative = "less", paired = TRUE)

##
## (3)
##

## Normal-based upper bound
mean(x) + qt(0.9, n - 1) * sd(x) / sqrt(n)

## Distribution-free upper bound
pbinom(1:n, n, 0.5)
sort(x)[17]

####
#### (5)
####

n <- 100
Y <- 92
p_hat <- Y / n

##
## (1)
##

## Agresti-Coull interval for p
Y_tilde <- Y + 0.5 + qnorm(0.975) ^ 2 / 2
n_tilde <- n + qnorm(0.975) ^ 2
p_tilde <- Y_tilde / n_tilde

p_tilde + c(-1, 1) * qnorm(0.975) * sqrt(p_tilde * (1 - p_tilde) / n_tilde)

##
## (2)
##

## Binomial test
p_value <- 1 - pbinom(Y - 1, n, 0.8)

## Check size of test. Need to increment the critical value by one to achieve size no
## more than 0.05. The new critical value is 87.
Y_crit <- qbinom(0.95, n, 0.8)
1 - pbinom(Y_crit - 1, n, 0.8)
1 - pbinom(Y_crit + 1 - 1, n, 0.8)

## Asymptotic test
test_stat <- sqrt(n) * (p_hat - 0.8) / sqrt(0.8 * 0.2)
p_value <- 1 - pnorm(test_stat)

##
## (3)
```

```
##

p_1 <- c(0.75, 0.80, 0.85, 0.90, 0.95)
gamma <- 1 - pbinom(87 - 1, 100, p_1)

##
## (4)
##

## Based on exact test. Taking n = 82 gives alpha <= 0.05 and gamma >= 0.08.
n_seq <- 50:150
k <- length(n_seq)
Y_crit_seq <- alpha_seq <- gamma_seq <- numeric(k)
for(i in 1:k) {
  Y_crit_seq[i] <- qbinom(0.95, n_seq[i], 0.8)
  alpha_seq[i] <- 1 - pbinom(Y_crit_seq[i] - 1, n_seq[i], 0.8)
  ## Increment critical value if alpha > 0.05
  if(alpha_seq[i] > 0.05) {
    Y_crit_seq[i] <- Y_crit_seq[i] + 1
    alpha_seq[i] <- 1 - pbinom(Y_crit_seq[i] - 1, n_seq[i], 0.8)
  }
  gamma_seq[i] <- 1 - pbinom(Y_crit_seq[i] - 1, n_seq[i], 0.9)
}

## Check size and power for selected n
Y_crit <- Y_crit_seq[n_seq == 82]
alpha <- 1 - pbinom(Y_crit - 1, 82, 0.8)
gamma <- 1 - pbinom(Y_crit - 1, 82, 0.9)

## Based on asymptotic test
((qnorm(0.95) * sqrt(0.8 * 0.2) + qnorm(0.8) * sqrt(0.9 * 0.1)) / 0.1) ^ 2

####
#### (MC4)
####

gamma <- 1 - pchisq((23.8 ^ 2 / 47.9 ^ 2) * qchisq(0.95, 9), 9)
1 - gamma

####
#### (MC5)
####

## 95th percentile of null Binomial distribution
qbinom(0.95, 20, 0.2)

## Level of test based on this critical value = 0.087. Adjust critical value, achieving
## level 0.032.
1 - pbinom(6, 20, 0.2)
1 - pbinom(7, 20, 0.2)

## Now compute power at p_1 = 0.4. Probability of Type II error is one minus this
1 - pbinom(7, 20, 0.4)
```

```
pbinom(7, 20, 0.4)

####
#### (MC6)
####

Delta <- sqrt(15) * 0.8
pt(qt(0.99, 14), 14, Delta)
```