k: # of subgroups is 25 n=3 ← Size of subgroup ①

1. I have more time in the summer for walking my dogs. In fact, I was able to take each of my three dogs on a separate walk every day. For 25 consecutive days, I recorded the amount of time that I walked each dog. The Minitab worksheet **Exam2_DATA_Thursday** contains these walk times in columns C2-C4.

(a) [+3] Create *Xbar* and *R* charts by averaging the times for the daily three walks. Using the control charts, estimate the process mean $\hat{\mu}$ and process standard deviation $\hat{\sigma}$. Show your work for computing $\hat{\sigma}$. Assume the process is in-control to compute these values. Report your answers correct to ~~3~~ 2 decimal places.

+3 total $\hat{\mu} =$ 27.78 +1 $\hat{\sigma} =$ 9.90/d_2 for $n=3$ $\frac{9.90}{1.693} \cong$ 5.848 +2 OR 5.85

(b) [+2] True or False. I performed Anderson-Darling normality tests for each dog's walk times. I obtained *p*-values that were all greater than 0.2. We can conclude that each dog's walk times are from a normally distributed population.

True +2 False

(c) [+3] From what we've discussed *in this class*, the AD normality test returns *p*-values less than the Ryan-Joiner *p*-values when the data …

A. Is trending up or down

B. Is rounded and appears to be integer data +3

C. Has a normally distributed histogram

~~D. Is from a small sample, such as $n < 15$~~

E. Is from a large sample, such as $n > 50$

(d) [+4] By the way that I subgrouped the data, the points on the ____________ chart indicate the spread or variation of walk times *between* dogs.

True Story!

Xbar +4 *Range* ✓

Can check: Day 1: 29.9 − 21.4 = 8.5 Day 2: 31.3 − 24.2 = 7.1

Note: I took Shay for shorter walks – he was much older than the other two

(e) [+3] Minitab is noting that a point is out-of-control on the *Xbar* chart at point 23. What is the reason why the point is considered out-of-control?

A. Two out of three consecutive points fall outside the 2σ warning limits on the *same side* of the center line

B. Four of five consecutive points fall beyond the 1σ limit on the *same side* of the center line Nope

C. Nine or more consecutive points fall to one side of center line Nope

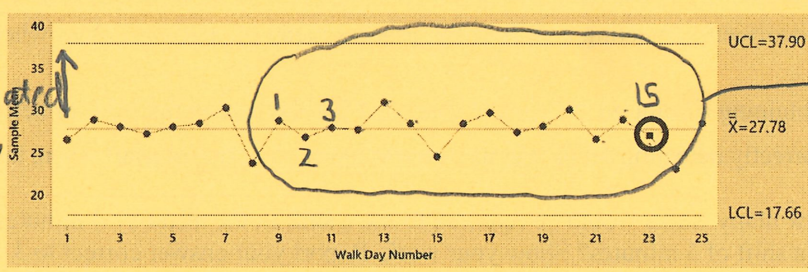
D. A run of six or more consecutive points increasing or decreasing. Nope

E. Fifteen points in a row within 1σ of center line (either side) Nope

+3

min time is usually Shay, the oldest dog

exaggerated UCL, LCL



1 2 3 15 string of "too good to be true" points

(f) [+3] Since the points on the *Xbar* chart are so "tight" along the centerline, this is an indication that all three dogs are getting approximately the same daily walk times, on average.

True False +3

Shay is always least time

(g) [+3] Since the points on the *Xbar* chart are so "tight" along the centerline and the control limits are exaggerated, this is an indication that there is more variation …

+3 A. Between the dogs' walking times than within each dog's walking times.

B. Within each dog's walking times than between the dogs' walking times.

It's tight because we are continually averaging a low, low, high

+21 ✓

- each dog's walking time is similar within that dog
- the walking times between the dogs is larger since Shay is always least time

If use $\hat{\mu} = 19.22$, LCL = 16.05, but let $\hat{\sigma} = 1$, will get 0.0007622 ②

#5 wrong answer

Problem 1 continued. For parts (h)-(m), we'll just consider Shay's walk times. Assume his walk times are ***normally distributed*** for the following parts.

(h) [+3] Create *I* and *MR* charts by using Shay's walk times. Using the control charts, estimate the process mean $\hat{\mu}$ and process standard deviation $\hat{\sigma}$. Show your work for computing $\hat{\sigma}$. Report your answers correct to ~~3~~ 2 decimal places.

$\hat{\mu} = 22.22$ +1

$\hat{\sigma} = \dfrac{2.321}{1.128} \approx 2.058$ +2 OR 2.06

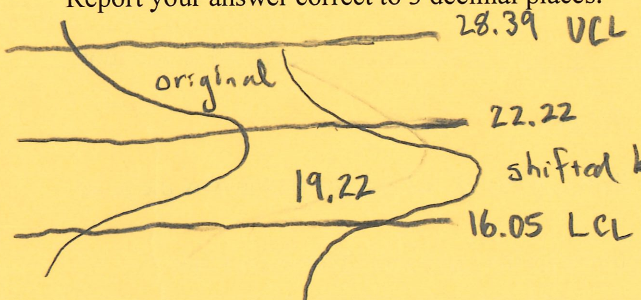
(i) [+3] What's the probability of committing a Type I Error on the *I-MR* chart? You do not need to show any work on this part. Report your answer correct to 3 decimal places.

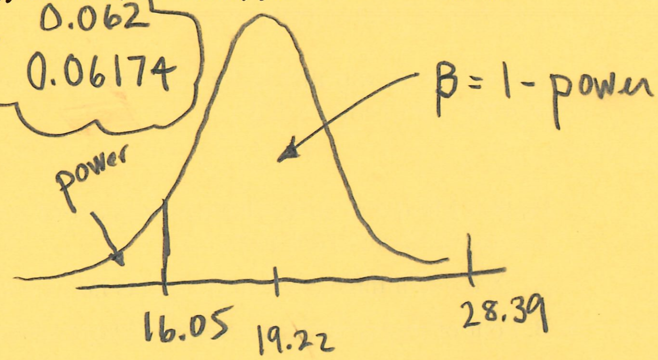
As long as UCL and LCL are set as 3 std devs above and below the centerline, the Type I error is: $2 \cdot (0.00135) =$ 0.0027 (0.0026) 0.003

for all I-MR, $\bar{X}$-R charts.

(j) [+5] As school time nears, my walk times ***get shorter***. Suppose my walk times with Shay shift by 3 minutes. What's the *power* associated with this shift on the *I-MR* chart? Show your work, which may just be a Minitab graphic. Report your answer correct to 3 decimal places.

0.062
0.06174





power: Prob Dist Plot
$\hat{\mu} = 19.22$, $\hat{\sigma} = 2.058$

For $\hat{\mu} = 19.22$, $\hat{\sigma} = 2.06$: 0.06192 still 0.062

(k) [+3] What is power with respect to a control chart?

A. A large, unbiased sample

B. The probability of correctly detecting a shift in the process mean on a control chart when, in fact, there is a shift.

C. A small amount of variability in the population

D. The probability of incorrectly detecting a shift in the process mean on a control chart when, in fact, there is no shift. Type I

E. The probability of obtaining a point on the control chart that is either above the UCL or below the LCL.

F. The probability of correctly detecting no shift in the process mean on a control chart when, in fact, there is no shift. no error

(l) [+3] On average, how many walks would be required to see a shift in the mean by 3 minutes? That is, what is the *average run length* for a shift of 3 minutes? Show your work. Report your answer correct to 3 decimal places.

$$\text{ARL} = \frac{1}{\text{power}} = \frac{1}{0.06174} \approx 16.207 \quad \text{OR} \quad \frac{1}{0.062} \approx 16.129$$

+2   +1 for correct value

Important: ARL = $\dfrac{1}{\text{answer from j}}$

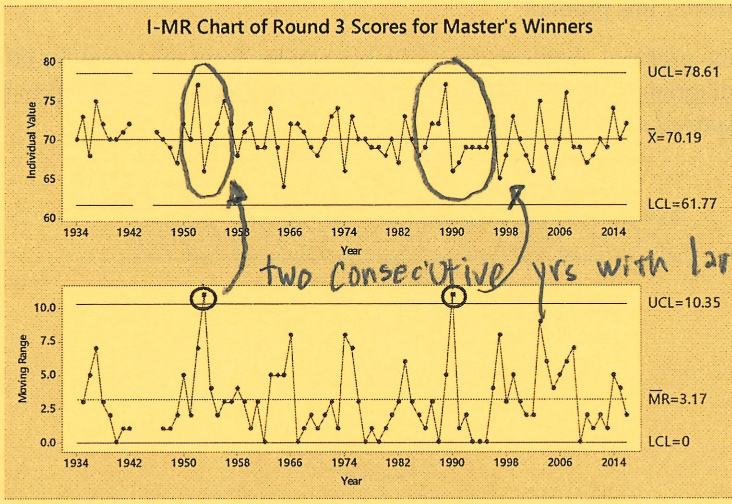
(m) [+2] How can we increase power without increasing Type I Error?

+2 increase sample size

+19

1(j) Issues - don't include both tails. -2; uses $\hat{\sigma}$ incorrect - divided by wrong d_2, used an incorrect formula, etc. Regraded exams w/ student's value for $\hat{\sigma}$

2. Consider the **Day 3 Scores** of the overall winners for the Masters Golf Tournament from years 1934 to 2015, where four years are missing due to World War II. Each point on the *I* chart below is the winning golfer's 3rd round score. The bottom chart is the *MR* moving range chart. The year, winning player, and score are in Minitab columns C6-C8.



two consec'utive yrs with large spread

(a) [+3] There are two out-of-control points circled above on the moving range chart (e.g., 1953 is one). What do these out-of-control points on the *MR* chart indicate about the winning golfers' scores for the 3rd round?

A. The winning golfer for that year had a large range for his 4 day golf scores. Range chart, not MR

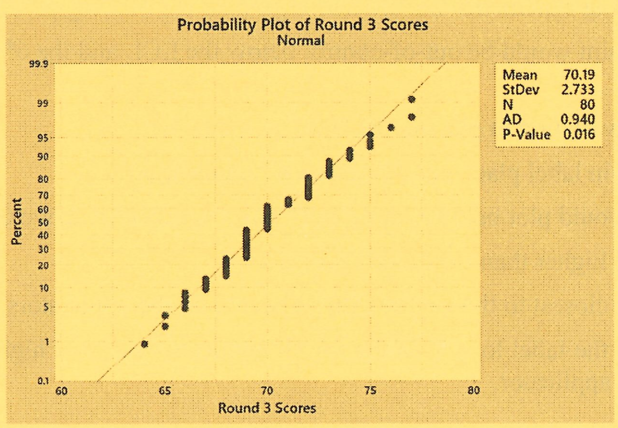
+3 B. There was a large difference between the winning golfer's score in one year compared to the winning golfer's score in the next year. MR-|difference in 2 consecutive pts|

C. The winning golfer had a large difference in his winning score compared to the 2nd place golfer's score for that year.

D. There was a large difference in the winning golfer's Day 3 score compared to his Day 3 scores from other golf tournaments besides the Masters.

E. It's the maximum score minus the minimum score for all golfers playing in the tournament that year.

(b) [+3] The following is a normality plot for the winners' Day 3 scores. Why is the Anderson-Darling *p*-value so low even though the data appears to be from a normal distribution?



RJ is > 0.10 because it doesn't penalize for data appearing integer-like OR discrete

+3 A. Because the golf scores are discrete values and there are "stacks" of points at integer values.

B. Because we forgot to add a decimal point and then a 0 after the scores to spread them out evenly along the plotted normality line. In other words, we should have 70.0 instead of 70, and 73.0 instead of 73, and 68.0 instead of 68, …

C. Because the golf scores do not follow a linear pattern.

D. Because the golf scores need to be transformed, and then they will appear normally distributed.

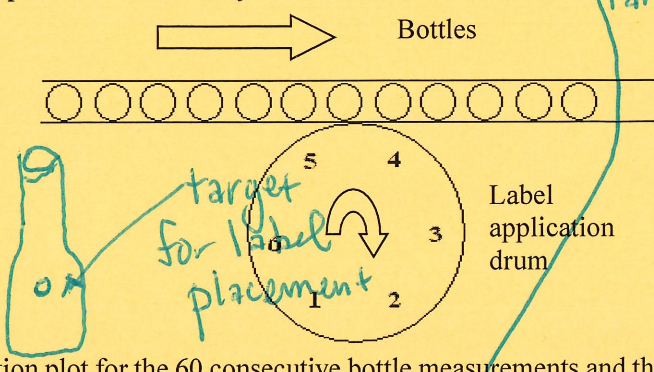
E. Because it was computed incorrectly in Minitab.

+6

3. True Story: According to a speaker from Frito Lay who spoke to my Quality Control class, Pepsico was having problems with Gatorade bottle logos on their containers. The problem occurred due to off-center labels on the bottles. The excessive variation in the position of the labels detracted from the appearance of the product and this was affecting the company's sales of this product.

To gain some insight into the problem, the deviations of label heights from the target for 60 consecutive bottles were measured. The schematic diagram in the figure below indicates how labels were applied to bottles by a *rotating drum* with *six label applicators spaced around its surface*.

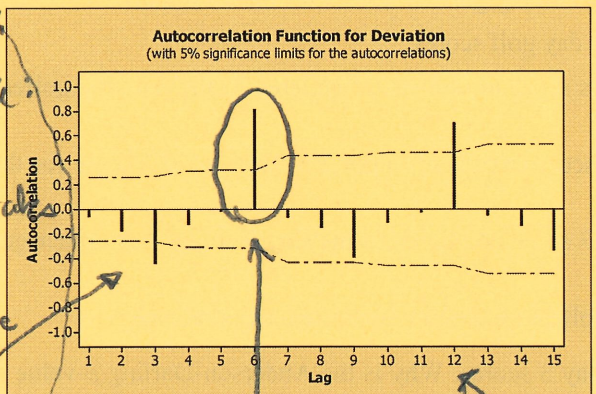
target is center placement on the bottle

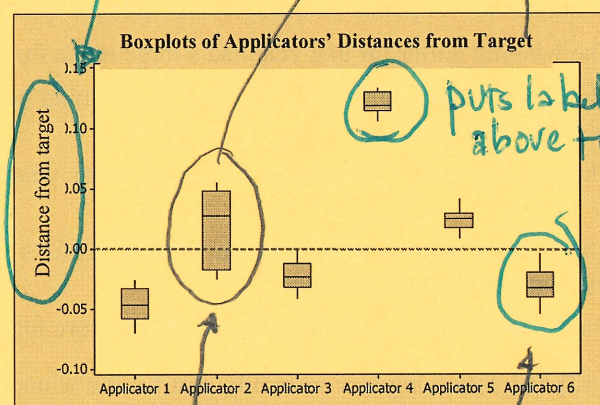


a lot of spread in Applicator 2's placement

Below are the autocorrelation plot for the 60 consecutive bottle measurements and the boxplots of deviations of label heights from the target for each of the six applicators.

Grading: One mistake: -1.5
two mistakes -3
3 or more -5





puts label WAY above target

puts label BELOW target

(a) [+5] Circle ALL the following statements that are true regarding the bottling process.

~~A.~~ On a control chart, the 3rd point would be out-of-control below the LCL and the 6th point would be out-of-control above the UCL. that's not what the lags mean!; also, on avg, 6th pt will not be above UCL

B. Applicator 2 has the largest variation in placing the label on the bottle.

C. There is positive correlation in label placement for every 6th application.

~~D.~~ On an *I* chart, the 3rd point would plot below 0 and the 6th point would plot above 0. not what lags at 3 and 6 mean

~~E.~~ Applicator 6 places the label higher than the other applicators do. boxplot doesn't suggest this

~~F.~~ Applicator 1 places the label closest to the center with respect to all the applicators. lag at 1 doesn't mean this; boxplot doesn't support

G. If the first applicator places the label low, then the fourth applicator places it high, and vice versa. There is this flip-flop pattern for every third applicator. Yep! that's what the negative lag at 3 means

(b) [+4] *Xbar-R* charts for tracking deviations from the target by subgrouping bottles *across* Applicators 1–6 would most likely produce an Xbar chart with the following characteristic.

+4 A. A "too good to be true" pattern with consecutive subgroups plotting close to the chart's centerline — lot of variation across applicators

B. A trending pattern, such as a string of 6 consecutive points increasing or 6 consecutive points decreasing

C. Points that plot outside the Xbar chart's lower or upper control limit

D. A jig-jag pattern (up and down pattern) that indicates tampering with the process

+9

#5. Using Largest Extreme Dist: LSL = 2, USL = 6
prob: ≈ 0.4488

4. [+4] A process cuts thin tubes of plastic into useable drinking straws for toddlers. The following samples of subgroups of size $n = 5$ are lengths (in mm) of straws drawn from the process when it was known to be "in-control." The lengths are assumed to be normally distributed.

LSL = 2, USL = 5, Prob ≈ 0.3542

Subgroup Number	Observation 1	Observation 2	Observation 3	Observation 4	Observation 5	Ranges
1	79.5	78.8	80.1	78.4	81.0	**2.6**
2	80.5	78.7	81.0	80.4	80.1	**2.3**
3	79.8	79.9	80.4	80.3	80.8	**1.0**
4	78.9	79.4	79.7	79.6	80.6	**1.7**
5	80.5	79.6	80.4	80.8	78.8	**2.0**
6	79.8	80.6	80.5	80.0	81.1	**1.3**

Note: $\bar{\bar{X}} = 80.0$ **mm and** $\bar{R} = 1.82$ **mm.**

What are the approximate upper control limits of the $\bar{X}$ and R charts for this process? Round calculated values to 2 decimal places. Use control chart constants with 3 decimal place accuracy.

A. $UCL_{\bar{X}} \approx \cancel{80.88}$, $UCL_R \approx \cancel{3.81}$

B. $UCL_{\bar{X}} \approx \cancel{86.01}$, $UCL_R \approx \cancel{8.81}$

+4 C. $UCL_{\bar{X}} \approx 81.05$, $UCL_R \approx 3.85$

+2 D. $UCL_{\bar{X}} \approx \cancel{82.34}$, $UCL_R \approx 3.85$ partial credit

E. $UCL_{\bar{X}} \approx \cancel{80.92}$, $UCL_R \approx \cancel{3.81}$

+2 F. $UCL_{\bar{X}} \approx 81.05$, $UCL_R \approx \cancel{3.81}$ partial credit

$$UCL_{\bar{X}}: \bar{\bar{X}} + 3\cdot\frac{\bar{R}/d_2}{\sqrt{5}} = 80.0 + 3\cdot\frac{1.82/2.326}{\sqrt{5}} \quad \text{OR} \quad \bar{\bar{X}} + A_2\bar{R}$$

$$80.0 + 0.577\cdot 1.82$$

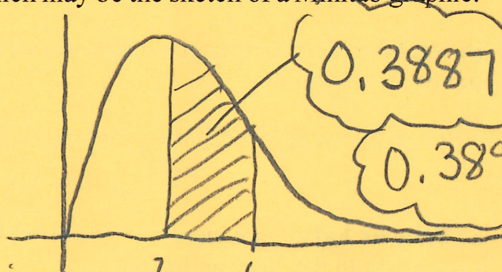
$$81.05$$

$$UCL_R = D_4\cdot\bar{R} = 2.115\cdot 1.82 \cong 3.85$$

5. [+6] We have a process that is *not* normally distributed. We use Minitab's "Individual Distribution Identification" function to suggest possible distributions that more accurately fit the process data. Using the most appropriate distribution from the Minitab output below, determine the probability that the data meets its specifications of LSL = 2 and USL = 6. Show your work below in determining this probability, which may be the sketch of a Minitab graphic. Report your answer correct to 3 decimal places.

most have changed it to USL = 5

```
Goodness of Fit Test
Distribution                 AD        P
Normal                    1.964   <0.005
Weibull                   0.183   >0.250
Smallest Extreme Value    4.032   <0.010
Largest Extreme Value     0.728    0.052
Logistic                  1.169   <0.005
```

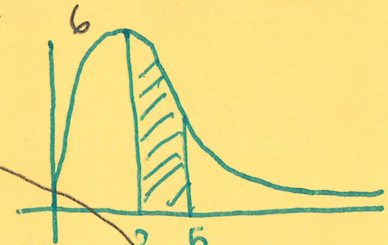


0.3887

0.389

If you used Largest Extreme Value, not the BEST choice SEE ABOVE

```
ML Estimates of Distribution Parameters
Distribution              Location    Shape     Scale  Threshold
Normal                     4.95512             4.80387
Weibull                              1.07049   5.09146
Smallest Extreme Value     7.67084             6.47704
Largest Extreme Value      3.02075             2.97033
Logistic                   4.24323             2.36915
```



Prob Dist Plot > Weibull w Shape Parameter and Scale Parameter >

Shaded Area: X value, Middle, X value 1 = 2
X value 2 = 6

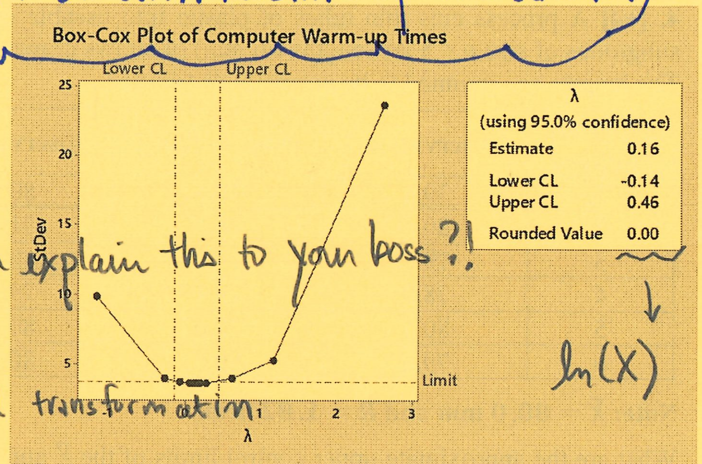
+10

some exams LSL = 2, USL = 5 : 0.3173

some exams LSL = 2, USL = 6 : 0.3887

When checking answers – notice the wording since I changed the T/F on diff. versions of the exam.

6. [+4] Computer warm-up times X will be plotted on an *I-MR* chart. Unfortunately, the warm-up times are not from a normal distribution and a transformation of the data is required to "convert" the times to normally distributed data. According to the Box-Cox plot and the fact that you'll need to explain the transformation to your manager, which is the most appropriate transformation to use?



+4 A. $\ln(X)$ OR +2 B. $X^{0.16}$ ← how do you explain this to your boss?!

↓ $\ln(X)$

C. X^1

D. $\sqrt{X}$

E. e^X

F. X^0 this is not a valid transformation

7. [+24, +3 each] True/False Statements There are multiple versions of the exam – these are T/F answers for in-class version

A. True or False. When selecting a rational subgroup for an *Xbar-R* chart, we do so in a way that *minimizes the variation within that subgroup* and *maximizes the opportunity for variation between subgroups.*

Exactly!

True False

B. True or False. A process can be in control on an *I-MR* chart, but not capable of satisfying a customer's specification requirements.

definitely eg 4 2 But specs: 7 to 9

UCL = 6 LCL = 2, But specs not even close to those eg LSL = 7 USL = 9

True False

C. The diameter of a bad brand of tennis balls has a non-normal distribution with a mean of $\mu = 2.58$ inches and a standard deviation of $\sigma = 0.05$ inch. Let X represent the diameter.

True or False. We can determine the following probability for one ball from this population: $P(X < 2.57)$.

No – we don't know which non-normal dist that we have

True False

True or False. We can determine the following probability for $n = 50$ balls from this population: $P(Xbar < 2.57)$.

Yes! By the CLT, $\bar{X}$ is normal for $n = 50$

True False

D. Suppose the diameter of the balls (from the problem above) is from a normal distribution with the same mean and standard deviation.

$\bar{X}$ ← tighter around 2.58 than X

True or False. $P(X < 2.57)$ would be smaller than $P(Xbar < 2.57)$. No work necessary.

True False

2.57 2.58 2.57

E. True or False. In general, the probability of committing a Type I error using only Rule 1 for special causes is the same on both the *Xbar* and *I* charts.

always $+3\sigma$ and -3σ beyond the CL

True False

F. True or False. When creating an *Xbar-R* control chart, using all the Minitab tests for special causes increases the probability of committing a Type II Error.

Type I error ↑

True False

G. True or False. On an *I-MR* chart, suppose the probabilities of committing Type I Errors by applying 3 different special cause rules are 0.05, 0.045, and 0.08, respectively. The overall probability of committing a Type I Error if all three rules are applied at the same time is approximately 0.195. Assume independence of the rules.

P(break at least one special cause rule) =

True False

+28

1 – P(break NO special cause rules) =

$1 - (0.95)(0.955)(0.92) \doteq 0.165$