AP Statistics Ch 6 - Extra Practice **Name:** Solutions **Per:** ____

Ninety-eight women and 225 men participated in a five kilometer road race. Below are summary statistics from Minitab on their times in the race.

Descriptive Statistics: Women, Men

Variable	N	Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Women	98	28.891	6.481	17.717	24.229	26.917	32.525	48.800
Men	225	24.758	5.026	15.050	21.117	24.050	27.608	48.417

#1. Santiago finished in 51st place among the men, with a time of 20.9 minutes. Keisha finished 33rd among women, with a time of 25.2 minutes. Use percentiles and z-scores to compare Santiago's and Keisha's relative standing among men and women, respectively.

Santiago

$\frac{51}{225} = .23$ (top 23%)

$1 - .23 = .77$ (77th percentile)

$z = \frac{20.9 - 24.758}{5.026} = -0.7676$

Keisha

$\frac{33}{98} = .34$ (top 34%)

$1 - .34 = .66$ (66th percentile)

$z = \frac{25.2 - 28.891}{6.481} = -0.5695$

Santiago finished better, faster than 77% of the men, compared to Keisha who was faster than 66% of the women. Santiago is 0.77 std dev below avg men's time compared to Keisha at 0.57 std dev below avg women's time.

#2. Suppose the timers of the race discovered that they accidently started the clock 15 seconds before the race actually started, so that each racer's finish time should be 15 seconds less. Describe the impact this would have on the mean, median, standard deviation, and interquartile range.

Mean: decrease by 15 seconds
median: decrease by 15 seconds
std dev: unaffected
IQR: unaffected

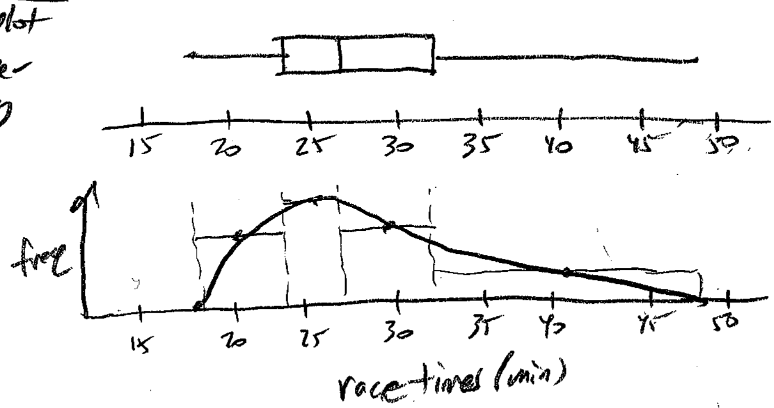
(add/subtract, shifting, does not affect measures of spread)

#3. Use the information in the computer output above to sketch a possible density curve for the finish times of women in this race. Be sure to provide a scale and label for the horizontal axis.

1) build a boxplot from 5 number summary

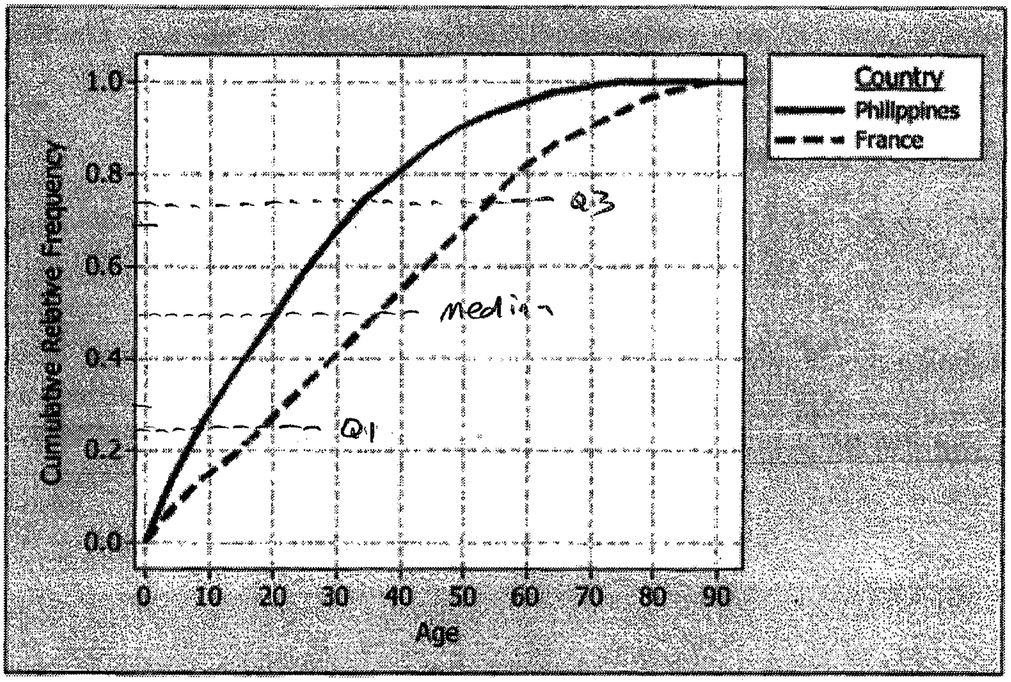
2) use 'aquarium' method.

3) smooth density curve



Below are cumulative frequency graphs for the age distributions of the populations of France and the Philippines.

this is an ogive (not a normal probability plot)



#4. Use the graphs to compare the median and interquartile range of ages in the two countries.

Philippines

Q1 = 8
med = 20 IQR = 34 − 8 = 26
Q3 = 34

France

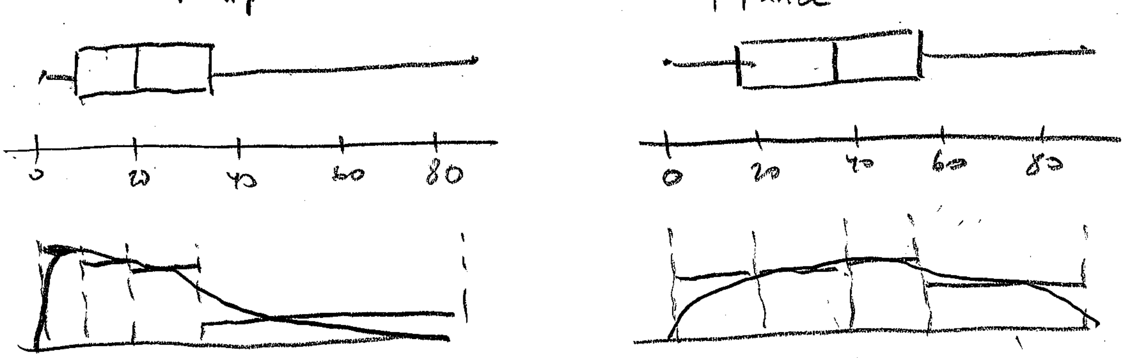
Q1 = 17 IQR = 55 − 17 = 38
med = 36
Q3 = 55

The median in France (36) is higher than in the Phillipines (20).
The IQR is larger in France (38) compared to the Phillipines (26).
People are generally older in France, ~~but there~~ is less consistency in ages in France also.

#5. One of these distributions is strongly skewed to the right. Which one is it, and how can you tell?

Phillipines

France



The Phillipines distribution is more strongly skewed right.

10. Car speeds. John Beale of Stanford, CA, recorded the speeds of cars driving past his house, where the speed limit read 20 mph. The mean of 100 readings was 23.84 mph, with a standard deviation of 3.56 mph. (He actually recorded every car for a two-month period. These are 100 representative readings.)

a) How many standard deviations from the mean would a car going under the speed limit be?

b) Which would be more unusual, a car traveling 34 mph or one going 10 mph?

a) $z_{20} = \frac{20 - 23.84}{3.56} = \boxed{-1.078}$

b) $z_{34} = \frac{34 - 23.84}{3.56} = +2.58$ The 10 mph car is more unusual.

$z_{10} = \frac{10 - 23.84}{3.56} = -3.88$

14. Caught speeding. Suppose police set up radar surveillance on the Stanford street described in Exercise 10. They handed out a large number of tickets to drivers going a mean of 28 mph, with a standard deviation of 2.4 mph, a maximum of 33 mph, and an IQR of 3.2 mph. Local law prescribes fines of \$100 plus \$10 per mile per hour over the 20 mph speed limit. For example, a driver convicted of going 25 mph would be fined 100 + 10(5) = \$150. Find the mean, maximum, standard deviation, and IQR of all the potential fines.

speed → mph over limit → fines

speed:
$\bar{x} = 28$ mph
$s = 2.4$ mph
(max = 33 mph
IQR = 3.2 mph)

−20

mph over limit:
$\bar{x} = 28 - 20 = 8$ mph
$s = 2.4$ mph
max = 33 − 20 = 13 mph
IQR = 3.2 mph
(spreads aren't affected)

×(\$10)
+(\$100)

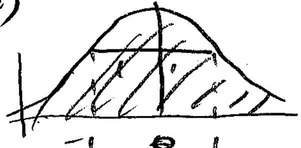
fines:
$\bar{x} = 8(10) + 100 = \$180$
$s = 2.4(10) = \$24$
max = 13(10) + 100 = \$230
IQR = 3.2(10) = \$32

(two transformations)

30. Normal models, again. What percent of a standard Normal model is found in each region? Draw a picture first.

a) $z > -2.05$
b) $z < -0.33$
c) $1.2 < z < 1.8$
d) $|z| < 1.28$

a)

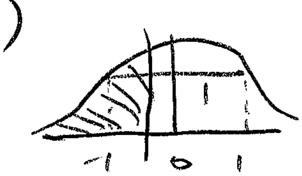


lower upper mean SD

normalcdf(-2.05, 999, 0, 1)

= .9798 $\boxed{(97.98\%)}$

b)

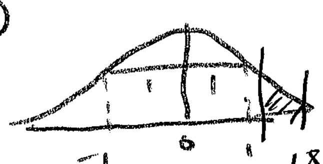


lower upper mean SD

normalcdf(-999, -.33, 0, 1)

= .3707 $\boxed{(37.07\%)}$

c)

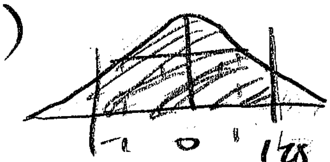


normalcdf(1.2, 1.8, 0, 1)

lower upper mean SD

= .0791 $\boxed{(7.91\%)}$

d)



lower upper mean SD

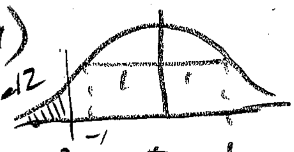
normalcdf(-1.28, 1.28, 0, 1)

= .7994 $\boxed{(79.94\%)}$

32. Yet another Normal model. In a standard Normal model, what value(s) of z cut(s) off the region described? Remember to draw a picture first.

a) the lowest 12%
b) the highest 30%
c) the highest 7%
d) the middle 50%

a)



area to left mean SD

invNorm(.12, 0, 1)

z = $\boxed{-1.175}$

b)

area to left mean SD

invNorm(.70, 0, 1)

z = $\boxed{.524}$

c)

area to left mean SD

invNorm(.93, 0, 1)

z = $\boxed{1.476}$

d) .25 .50 .25

area to left mean SD

invNorm(.25, 0, 1)

$\boxed{z = -.674 \text{ to } +.674}$

40. Parameters II. Every Normal model is defined by its parameters, the mean and the standard deviation. For each model described below, find the missing parameter. Don't forget to draw a picture.

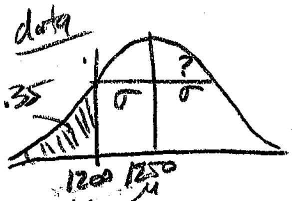
a) $\mu = 1250$, 35% below 1200; $\sigma = ?$

b) $\mu = 0.64$, 12% above 0.70; $\sigma = ?$

c) $\sigma = 0.5$, 90% above 10.0; $\mu = ?$

d) $\sigma = 220$, 3% below 202; $\mu = ?$

a) data



z-score

area to left, mean, SD

invNorm(.35, 0, 1)

$z = -.38536$

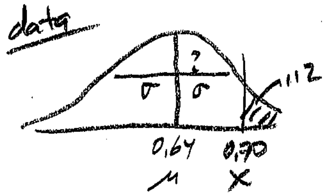
$z = \frac{x - \mu}{\sigma}$

$-.38536 = \frac{1200 - 1250}{\sigma}$

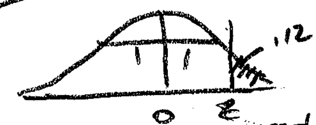
$\sigma = \frac{1200 - 1250}{-.38536}$

$= \boxed{129.76}$

b) data



z-score



area to left, mean, SD

invNorm(.88, 0, 1)

$z = 1.17498$

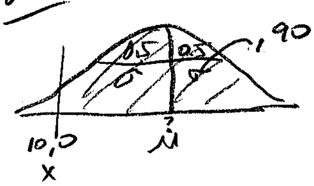
$z = \frac{x - \mu}{\sigma}$

$1.17498 = \frac{0.70 - 0.64}{\sigma}$

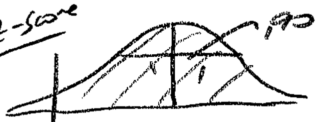
$\sigma = \frac{0.70 - 0.64}{1.17498}$

$= \boxed{0.051}$

c) data



z-score



z, area to left, mean, SD

invNorm(.10, 0, 1)

$z = -1.28155$

$z = \frac{x - \mu}{\sigma}$

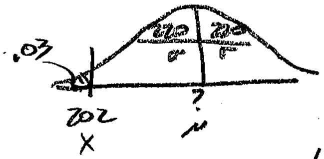
$-1.28155 = \frac{10.0 - \mu}{0.5}$

$(-1.28155)(0.5) = 10 - \mu$

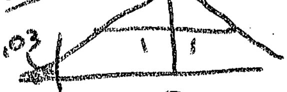
$\mu = 10 + (1.28155)(.5)$

$= \boxed{10.64}$

d) data



z-score



z, area to left, mean, SD

invNorm(0.03, 0, 1)

$z = -1.88079$

$z = \frac{x - \mu}{\sigma}$

$-1.88079 = \frac{202 - \mu}{220}$

$(-1.88079)220 = 202 - \mu$

$\mu = 202 + 220(1.88079)$

$= \boxed{615.77}$

49. Eggs. Hens usually begin laying eggs when they are about 6 months old. Young hens tend to lay smaller eggs, often weighing less than the desired minimum weight of 54 grams.

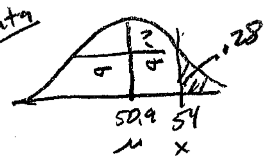
a) The average weight of the eggs produced by the young hens is 50.9 grams, and only 28% of their eggs exceed the desired minimum weight. If a Normal model is appropriate, what would the standard deviation of the egg weights be?

b) By the time these hens have reached the age of 1 year, the eggs they produce average 67.1 grams, and 98% of them are above the minimum weight. What is the standard deviation for the appropriate Normal model for these older hens?

c) Are egg sizes more consistent for the younger hens or the older ones? Explain.

d) A certain poultry farmer finds that 8% of his eggs are underweight and that 12% weigh over 70 grams. Estimate the mean and standard deviation of his eggs.

a) data



$z = \frac{x-\mu}{\sigma}$

$.58284 = \frac{54-50.9}{\sigma}$

$\sigma = \frac{54-50.9}{.58284}$

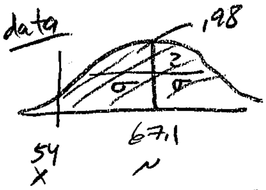
$= \boxed{5.3g}$

z-score

invNorm(.72, 0, 1) area to left, mean, SD

$z = .58284$

b) data



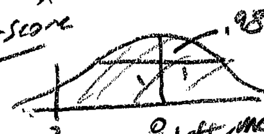
$z = \frac{x-\mu}{\sigma}$

$-2.05375 = \frac{54-67.1}{\sigma}$

$\sigma = \frac{54-67.1}{-2.05375}$

$= \boxed{6.4g}$

z-score

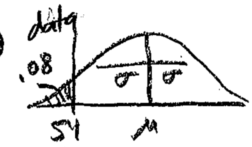


invNorm(.02, 0, 1) area to left, mean, SD

$z = -2.05375$

c) Egg sizes are more consistent for $\boxed{\text{younger}}$ hens, because the standard deviation is smaller.

d) data



z

z = invNorm(.08, 0, 1) area to left, mean, SD

$= -1.40507$

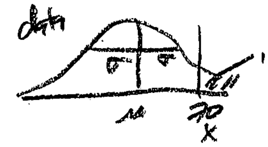
$z = \frac{x-\mu}{\sigma}$

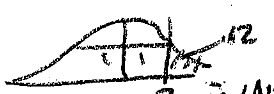
$-1.40507 = \frac{54-\mu}{\sigma}$

$-1.405070\sigma = 54 - \mu$

$\mu - 1.405070\sigma = 54$

data





z = invNorm(.88, 0, 1) area to left, mean, SD

$z = 1.174987$

$z = \frac{x-\mu}{\sigma}$

$1.174987 = \frac{70-\mu}{\sigma}$

$1.174987\sigma = 70 - \mu$

$\mu + 1.174987\sigma = 70$

a system of 2 equations, w/ 2 unknowns

$$\begin{cases} \mu - 1.405070\sigma = 54 \\ \mu + 1.174987\sigma = 70 \end{cases}$$

matrix form:

$$\begin{array}{ccc} \mu & \sigma & RHS \\ \end{array} \left[\begin{array}{cc|c} 1 & -1.40507 & 54 \\ 1 & 1.174987 & 70 \end{array}\right]$$

rref $\rightarrow$

$$\begin{array}{ccc} \mu & \sigma & RHS \\ \end{array} \left[\begin{array}{cc|c} 1 & 0 & 62.7 \\ 0 & 1 & 6.2 \end{array}\right] \rightarrow \boxed{\begin{array}{l} \mu = 62.7g \\ \sigma = 6.2g \end{array}}$$