Biostatistics Exam 01 Name:

1. Given these tabular data:

f_i is the frequency of occurrence of the value of X_i.

i	X_i	f_i	$f_i * X_i$
1	22	3	66
2	26	5	130
3	28	6	168
4	31	12	372
5	[redacted]		[redacted]
6	38	5	190
7	44	6	264

$\bar{X} = \frac{\sum x_i}{n}$

$\sum = 66 + 130 + 168 + 372 + 190 + 244 + Z =$

$= 1170 + Z$

$\bar{X} = 32.49$

$n = 45$

What does $f_5 * X_5$ equal? **(2)**

$32.49 = \frac{1170 + Z}{45}$

$(32.49)45 = 1170 + Z$

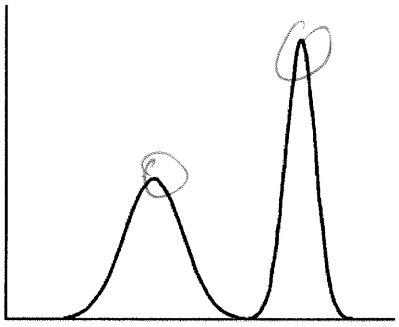
$Z = (32.49)45 - 1170$

$= 292.05$

Page total pts:

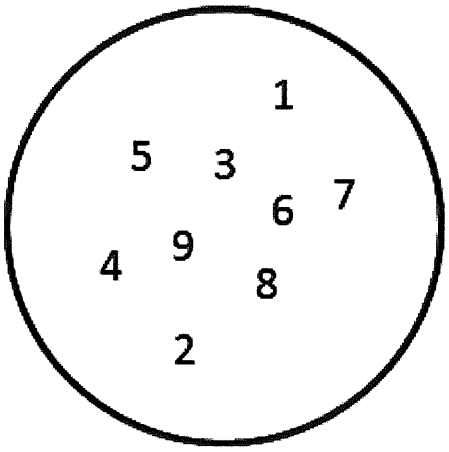
Biostatistics Exam 01 Name:

2. What is the most efficient measure of central tendency in this distribution, why? **(1)**



Mode, because it is bimodal — median an mean will not describe the central tendency — most freq. occuring ↓ hard question — Dr. Wu and I discussed — all get full credit!

3. Given this sample set:



What is the probability of drawing a number from the set [1,2,3] or a number from the set [8,9] **(1)**

$$P(1,2,3) = \frac{3}{9}$$

$$P(8,9) = \frac{2}{9}$$

$$P(A \cup B) = P(A) + P(B)$$

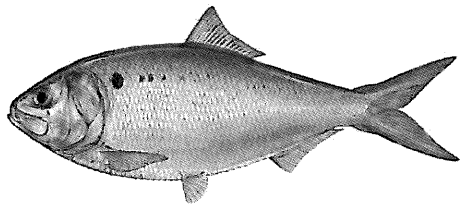
$$= \frac{3}{9} + \frac{2}{9}$$

$$= 5/9$$

Page total pts:

Biostatistics Exam 01 Name:

I am studying the lengths of fish in population of age-1 Gulf Menhaden, *Brevoortia patronus*. I have made a census of the population and measured every single fish ($n = 2.9 \times 10^{10}$) in the population. I have found that the distribution of lengths of fish can be described using a normal distribution. The mean value of the population is 100 mm and the population variance is 9 mm^2.



4. What is the proportion of the population that is less than 120 mm? **(1)**

< 120

range(rnorm(10,000, 100, 3)) = [~~65~~, ~~133~~]

So ~~zero~~ prob. of getting > 120... in this draw,

So prop ≈ 1

5. What is the proportion of the population that is greater than 105mm? **(1)**

1 − pnorm(105, 100, 3) = 0.0477

or 4.8 %

6. I take a sample a single randomly selected fish from the population. What is the probability of selecting an individual that is less than 120 mm in length and greater than 150 mm in length? **(1)**

P(<120) ∪ P(>150) = ∅ not possible empty set

∴ Pr. is zero

7. I have estimated the population parameters accurately. Is it possible to increase the precision of the estimate of the mean? If so, how and if not, why **(2) Your answer will begin with "Yes, it is possible" or "No, it is not possible".**

$\sigma^2 = 9\ mm^2$ is process error.

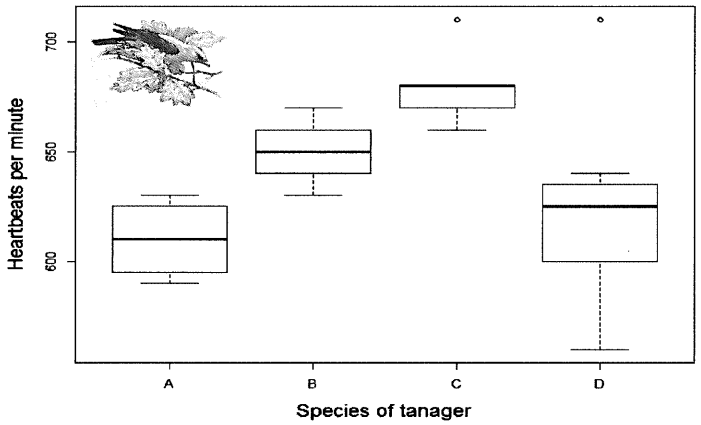
We measured every fish already....

no poss to increase precision

Page total pts:

Biostatistics Exam 01 Name:

I am interested in studying the physiology of tanager species in the northeast United States. I have done some preliminary data analysis, collected using random sampling, and created a boxplot describes the data that I have collected. The "box" is the value at mean ± one standard deviation, the thick line is the median, outliers are points, and the "caps" describe the mean ± two standard deviations.



8. Do of these species-specific measurements exhibit skewed distributions **(1)**?

C, D

9. Do of these species-specific measurements exhibit normal distributions **(1)**?

A, B

10. I have 10% more samples to allocate (10% of my original number of samples) – that is I have the money, time, and inclination to take some additional samples. What strategy would you use to decide how these samples should be allocated? **(2) Your answer will begin with "The specific strategy I will use is:"**

C, D needs more samples.
outliers & skew can be tightened up -
by can be a result of
sampling error

Page total pts:

Biostatistics Exam 01 Name:

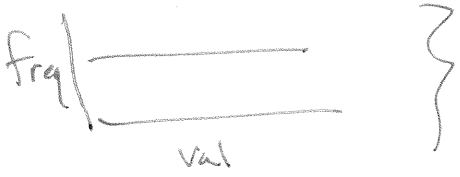
11. In a set of numbers that is composed of 100 elements in increasing order:

[1*λ, 2*λ, 2*λ, 4*λ, 5*λ, 6*λ .... 100*λ], where λ is positive and constant, what are the 20% and 80% quantiles? **(1)**

$[20\lambda, 80\lambda]$ let $\lambda = 1$.

12. Are the frequency of these data best modeled with a normal distribution? (why or why not) **(1)** **Your answer will begin with "Yes, it is best modeled with normal dist." Or "No, the normal model is not appropriate"**

No



not normal Equal prob for all outcomes

13. What is a logarithm **(1)**, you will need the word "exponent" in your explanation. I would like a definition, not an example.

the log of a number is the exponent that the base must be raised to.

14. What is the null hypothesis that you would evaluate for investigating whether the toxicity of fish from the Fukushima nuclear accident? **(1)** Fields p. 28

Spatial temporal: H_0: The is no difference in toxicity of fish sampled before & after nuclear accident

Spatial: H_0: There is no differe in toxicity of fish taken near or far from accident site