

STATISTICS 1040
Test 1, Spring 2012

Name _____
Recitation Instructor _____ Time _____

1. A recent issue of the New York Times included the following article:

Two widely used nutritional supplements for arthritis pain do not effectively relieve patients' aching arthritic knees, a large federal study has found. The study of two supplements, glucosamine and chondroitin sulfate, was published in The New England Journal of Medicine. The study was meant to provide a definitive answer as to whether these supplement work. In the 24-week study, 1583 patients with osteoarthritis of the knee were randomly assigned to one of five groups. Some patients took chondroitin and some took both. Others, serving as comparison subjects, took a placebo.

a) Is this an observational study? Briefly explain. (3 points)

No, the investigators decided who would be in the treatment group (randomly).

b) Why do they randomly assign the patients to the groups? (4 points)

Randomization makes the groups as much alike as possible and reduces bias.

c) What is a placebo and why is it used? (4 points)

A placebo is neutral but resembles the treatment. The response should be to the treatment itself and not the idea of treatment.

2. In 2006, a researcher claimed that elderly people who regularly indulged in "recreational gambling" were healthier, on average, than people who did not gamble at all. Provide a confounding factor that sheds doubt on the claim. (10 points)

Wealth could be a confounding factor. Wealthy people tend to be healthier and can also afford to gamble, to take trips to Vegas and Atlantic City.

3. Jeff took measurements of his diastolic blood pressure each day for a week. The values obtained were: 102, 95, 108, 103, 92, 97, and 117.

a) Compute the average and median of these values. (3 points)

$$AV = 102, \text{ Median} = 102$$

Compute the SD of the 7 measurements. (3 points)

$$SD = 7.89$$

b) Kristin also measured her diastolic blood pressure once a day for a week. Her measurements averaged 64 with an Sd of 8.2. If we combined our two sets of measurements, what would the SD of the combined group be equal to? (3 points)

i) about 8.2

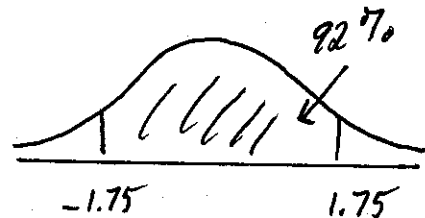
ii) less than 8.2

iii) more than 8.2

4. Many universities (including USU) require students to take the *Graduate Records Examination (GRE)* before being admitted to graduate school. The *Verbal GRE* is measured on a scale that runs from 200 to 800 points. Over the last 4 years the average Verbal GRE score was 460, the standard deviation was 120, and the histogram of Verbal GRE scores looked like the normal curve.

a. Approximately what percentage of students got 670 or higher on the Verbal GRE? (12 points)

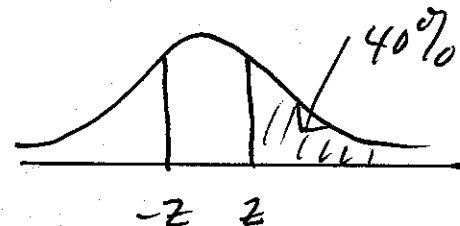
$$\frac{670 - 460}{120} = 1.75 \quad A(1.75) = 92\%$$



$$\frac{100 - 92}{2} = 4 \quad \boxed{4\%}$$

b. The Graduate School at a large state university requires students to score above the 60th percentile on the Verbal GRE. What is the score that represents the 60th percentile? (10 points)

$$A(z) = 20\% \quad \text{so } z = .25 \quad \frac{\text{score} - 460}{120} = .25$$

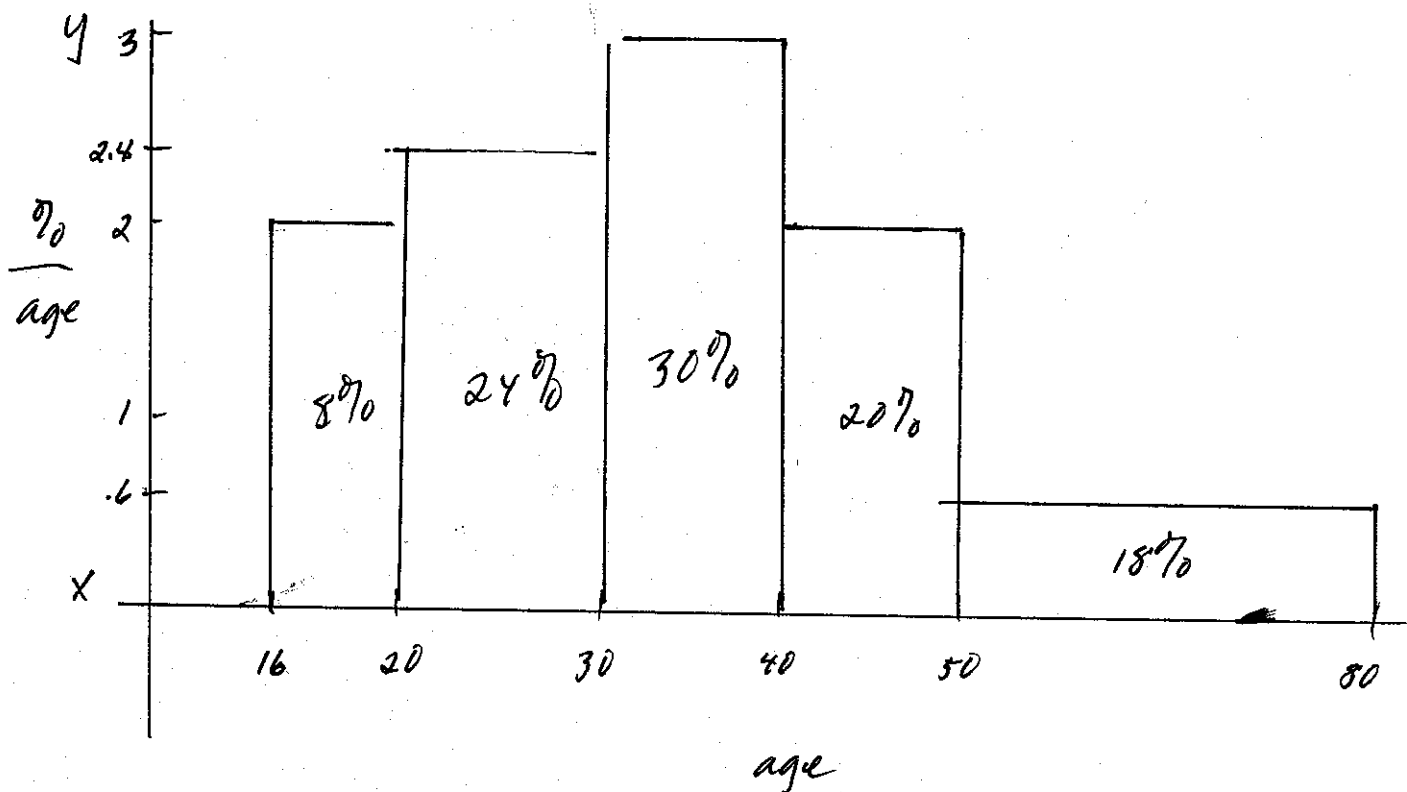


$$\text{score} = (120)(.25) + 460 = \boxed{490}$$

5. A study in a neighboring state showed that 4,750 drivers were stopped by police for a variety of reasons. The following table gives the ages of these drivers. Age intervals include the left endpoint but not the right.

Age	Number	%	Width	Height
16 - 20	380	8 %	4	2
20 - 30	1140	24%	10	2.4
30 - 40	1425	30%	10	3
40 - 50	950	20%	10	2
50 - 80	855	18%	30	.6

a) Draw the histogram. (15 points)



b) Estimate the percentage of drivers who were 35 years or older and stopped by the police. (3 points)

$$15\% + 20\% + 18\% = \boxed{53\%}$$

6. More than 5000 high-school students in California recently completed the SAT exam. The verbal SAT scores follow the normal curve with an average of 600 and a standard deviation of 100. The Math SAT scores for these students also had an average of 600 and an SD of 100. The correlation coefficient was .7 and the scatter diagram was football-shaped.

- a) Find the equation of the regression line for predicting the math SAT from the verbal SAT. (10 points)

$$\text{slope} = r \cdot \frac{SD_y}{SD_x} = (.7) \frac{100}{100} \cdot .7, \text{ point} = (600, 600)$$

$$y - 600 = .7(x - 600) = .7x - 420, \quad y = .7x + 180$$

- b) Use the regression method to predict the math SAT score for a student who scored 750 on the verbal SAT. (10 points)

Use the equation above or the 6-step method.

1. Verbal score
2. 750
3. $\frac{750 - 600}{100} = 1.5$
4. $.7(1.5)$
5. $(100)(.7)(1.5) = 105$
6. $600 + 105 = \boxed{705}$

- c) The r.m.s. error for predicting the math SAT from the verbal SAT is equal to $\sqrt{1 - r^2} \times SD_y = \sqrt{1 - (.7)^2} \times 100 \approx 71$. What does the r.m.s. error measure? (5 points)

It measures the overall size of the error in regression.

- d) If a student scored at the 80th percentile on the verbal SAT, do you expect them to score at about the 80th percentile on the math SAT? Explain. (5 points)

No, you would expect the student to score less than the 80th percentile because of the regression effect.