

1. An investor is rating two potential additions to a stock portfolio on the basis of the likely value of the stock after holding it one year. For each dollar invested, the person models

Stock A price after one year as normal with mean  $\mu = 1.15$  and  $\sigma = .05$

Stock B price after one year as normal with mean  $\mu = 1.25$  and  $\sigma = .10$

a) Which of the two stocks does this person judge most likely to be worth at least \$1.10 per dollar invested after one year? Show your work.

Circle one: Stock A or Stock B

b) The investor is 90% sure that Stock A will be worth *at least how much* per dollar invested after holding it one year?

2. The book *Data, Statistics and Decision Models with Excel* by Harnett and Horrell contains information from a real bill sent by the IRS to a real company demanding payment of back excise taxes. A (random) sample of  $n = 160$  company sales invoices (from a much larger pool of 22,794 invoices that were in dispute) produced  $\bar{x} = \$11.21$  owed per invoice with  $s = \$20.59$ .

a) We will assume that \$0 is the minimum that can be owed on any particular invoice. From this and the sketchy information about the sample supplied above, is it possible that the data set had a roughly normal/bell-shaped histogram? Why or why not?

b) Give approximate 90% confidence limits for the mean excise tax owed *per invoice*.

c) Does the validity of your answer in b) depend upon the 22,794 invoices having a normal relative frequency distribution of excise tax owed? Explain.

d) Apparently, the bill sent to the company was for \$194,647, or about \$8.54 per invoice. Put yourself in the position of an IRS supervisor. The sample mean (\$11.21) is substantially larger than the amount billed. Should you be convinced that the \$8.54 was clearly too low (and that your agents should clearly have sent a larger bill)? (Show appropriate null and alternative hypotheses and report an appropriate  $p$ -value.)

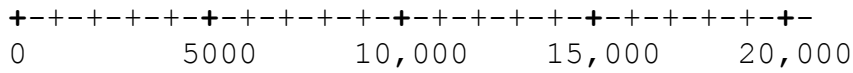
3. Attached to this exam is a JMP report for the analysis of some data from  $n = 65$  case files from an income tax preparation company. Represented are values of

$y =$  deductions claimed

and also  $\sqrt{y}$ .

a) Describe the shape of the  $y$  distribution in 10 words or less.

b) Using the scale below, make a boxplot for the deduction data.



c) Ignore the graduated nature of the income tax and suppose that each deduction dollar produces \$.28 worth of reduction in tax owed. That is, consider the quantity

$$x = .28y$$

What are the mean and standard deviation of  $x$ ?

$$\bar{x} = \underline{\hspace{2cm}} \qquad s_x = \underline{\hspace{2cm}}$$

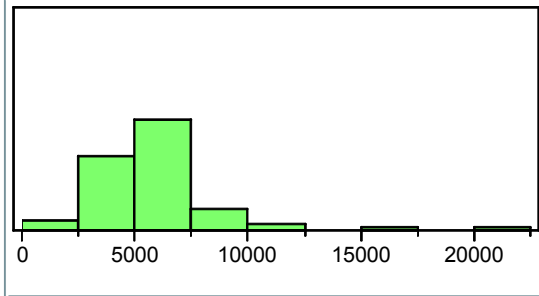
d) The first histogram suggests that using the raw deduction data,  $y$ , one should probably NOT use the formula from class to make a prediction interval for  $y_{\text{new}}$ . But in this regard, the second histogram looks "better" than the first. Use the second part of the JMP report and make a 95% prediction interval for  $(\sqrt{y})_{\text{new}}$ . Then square the end points of your interval to get a prediction interval for a single additional deduction value.

interval for an additional  $\sqrt{y}$  :

interval for an additional  $y$  from the above:

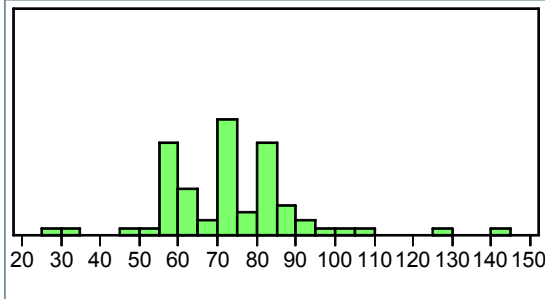
Distributions

DEDUCTS



| Quantiles |          | Moments |                |          |
|-----------|----------|---------|----------------|----------|
| 100.0%    | maximum  | 20701   | Mean           | 5791.385 |
| 99.5%     |          | 20701   | Std Dev        | 3097.026 |
| 97.5%     |          | 17844   | Std Err Mean   | 384.139  |
| 90.0%     |          | 8402    | upper 95% Mean | 6558.791 |
| 75.0%     | quartile | 6750    | lower 95% Mean | 5023.978 |
| 50.0%     | median   | 5450    | N              | 65.000   |
| 25.0%     | quartile | 3650    |                |          |
| 10.0%     |          | 3250    |                |          |
| 2.5%      |          | 938     |                |          |
| 0.5%      |          | 888     |                |          |
| 0.0%      | minimum  | 888     |                |          |

Root(DEDUCTS)



| Quantiles |          | Moments |                |          |
|-----------|----------|---------|----------------|----------|
| 100.0%    | maximum  | 143.88  | Mean           | 73.87286 |
| 99.5%     |          | 143.88  | Std Dev        | 18.42299 |
| 97.5%     |          | 133.36  | Std Err Mean   | 2.28509  |
| 90.0%     |          | 91.66   | upper 95% Mean | 78.43786 |
| 75.0%     | quartile | 82.16   | lower 95% Mean | 69.30786 |
| 50.0%     | median   | 73.82   | N              | 65.00000 |
| 25.0%     | quartile | 60.32   |                |          |
| 10.0%     |          | 57.01   |                |          |
| 2.5%      |          | 30.62   |                |          |
| 0.5%      |          | 29.80   |                |          |
| 0.0%      | minimum  | 29.80   |                |          |