Stat 2153 - Statistical Methods

Final - 2009.05.13

Name: ______________________________

Standard Deviation for a Set of Data:

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{n\sum(x^2) - \left(\sum x\right)^2}{n(n-1)}}$$

$$\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}}$$

Z-score for a Set of Data:

$$z = \frac{x-\bar{x}}{s} \text{ or } z = \frac{x-\mu}{\sigma}$$

Combination and permutation formulas:

$$_nP_r = \frac{n!}{(n-r)!}$$

$$_nC_r = \frac{n!}{(n-r)!r!}$$

Discrete Probability Distribution information:

$$\mu = \sum(x \cdot P(x))$$

$$\sigma^2 = \sum\left[(x-\mu)^2 \cdot P(x)\right]$$

$$\sigma = \sqrt{\sum\left[(x-\mu)^2 \cdot P(x)\right]}$$

Binomial Probability formula:

$$P(x) = \frac{n!}{(n-x)!x!} \cdot p^x \cdot q^{n-x}$$

Binomial Probability Distribution information:

$n =$ number of trials

$p =$ probability of success

$q =$ probability of failure

$$\mu = np$$

$$\sigma^2 = npq$$

$$\sigma = \sqrt{npq}$$

Central Limit Theorem and sampling distribution of $\bar{x}$:

If x has a normal distribution, or if the sample is large enough, with mean μ and standard deviation σ, then for a group of objects from the population, we have the following:

$$\mu_{\bar{x}} = \mu, \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Confidence Levels and Critical Values:

Confidence Level	α	Critical Value $z_{\alpha/2}$
90 %	0.10	1.645
95 %	0.05	1.960
99 %	0.01	2.575

Estimating a Population Proportion - Margin of error:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Estimating a Population Mean with σ Known - Margin of error:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Estimating a Population Mean with σ Not Known - Margin of error:

$$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Chi-Square Distribution:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

Test Statistic for Testing a Claim About a Proportion:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Test Statistic for Testing a Claim About a Population Mean with σ Known:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

Test Statistic for Testing a Claim About a Population Mean with σ Not Known:

$$t = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{s}{\sqrt{n}}}$$

Table 1: Days Until Tomato Plants Produce Fruit

65	67	70	72	74	75	76	76	76	77
78	79	79	79	80	80	80	80	80	81
82	84	85	86	87	88	88	89	90	91
91	94	96	98	99	100	101	104	105	106

1. Construct a frequency distribution table with class widths of 5 days, the first class having lower class limit 65 days.

2. Construct a histogram from the frequency distribution table.

3. Construct a frequency polygon using the classes defined previous.

4. Compute the mean, median and mode of the data. Also compute the mean from the frequency distribution table.

5. Compute the standard deviation of the data.

6. Determine Q_1, Q_2 and Q_3.

7. State the interquartile range for this data.

8. Point out any outliers in the data.

9. Find the z-score corresponding to the value of 89 days.

10. Construct a boxplot for the data in Table 1.

Table 2: Color and Type of Fruit

	tomato	hot pepper	sweet pepper
red	5	7	3
green	4	6	2
yellow	8	2	3
other	16	4	0

11. What is the probability that a single plant, picked at random, will be a tomato plant bearing green fruit?

12. What is the probability that a single plant, picked at random, is a pepper plant (hot or sweet) with red or green fruit?

13. What is the probability that a single plant, picked at random, will *NOT* be a tomato plant bearing yellow fruit?

14. If picking two plants without replacement, what is the probability that the first plant will be a hot pepper, with the second being a sweet pepper plant which bears fruit of a color other than yellow?

15. If picking three plants without replacement, what is the probability that the first will bear red fruit, the second green and the third yellow?

The following table shows the probabilities of the number of different types of tomatoes that will be produced over the course of a growing season.

number of types	probability
0	0+
1	0.001
2	0.002
3	0.014
4	0.022
5	0.033
6	0.041
7	0.078
8	0.102
9	0.153
10	0.230
11	0.193
12	0.131

16. Verify that the above table represents a probability distribution.

17. Compute the mean and standard deviation of the distribution.

18. Find the probability that at least six different varieties of tomatoes will be produced.

19. Find the probability that at most five different varieties of tomatoes will be produced.

Over the years, Karl has determined that a tomato plant has a 68% chance of producing at least 8 pounds of tomatoes over the growing season.

20. What is the probability that at least 22 out of 24 tomato plants will produce at least 8 pounds of tomatoes?

21. Planning for the future, it would be nice to know the probability of at least 190 out of 250 plants producing at least 8 pounds of tomatoes. Calculate this probability.

For problems 22 and 23, assume that the each tomato plant from last year yielded a mean weight of 8 pounds and standard deviation of 0.24 pounds of tomatoes. You may assume that the weights were normally distributed.

22. Find the probability that a randomly selected tomato plant will bear at least 8.43 pounds of tomatoes.

23. Find the probability that a selected plant will produce between 7.24 and 8.23 ounces of tomatoes.

24. A sample mean of 7.4 pounds was computed by picking fruit from 40 tomato plants, picked at random.. Construct a 95% confidence interval for the sample mean using the population standard deviation of 0.84 pounds. Does the population mean of 8 pounds lie in the confidence interval?

25. In problem 24, a sample mean of 7.4 pounds was computed by picking fruit from 40 tomato plants, picked at random. Test the hypothesis that the mean weight of the tomatoes off of each plant is equal to the population mean of 8 pounds. Once again, the population standard deviation is given to be 0.84 pounds.

a) What is the original claim?

b) State the values of n, $\bar{x}$, $\mu_{\bar{x}}$ and σ.

c) Determine H_0 and H_1.

d) Compute the test statistic for the sample mean.

e) Which tail-test is this?

f) Assuming a confidence level of 95%, what is the critical $z-$value?

g) Find the P-value for the test statistic from part d).

h) Do you reject the null hypothesis?

i) In a complete sentence, what can you conclude about the claim?