/40

STAT 2509 A

Assignment#1

SOLUTION

[3] **1.** a) Applied statistics can be divided into *descriptive statistics* and *inferential statistics*. Define the 2 statistics mentioned above.

Descriptive Statistics: methods for organizing, displaying and describing a group of data using graphs, tables and/or numerical measures (1)

Inferential statistics: methods for drawing conclusions and making inferences about the population based on a sample which was taken from our population of interest (1)

b) The process of using information from a sample to draw conclusions about the entire population is called

(i) sampling (ii) the scientific method
(iii) statistical inference (iv) descriptive statistics

(iii) statistical inference (1)

[1] **2.** Which of the following are measures of Variability?

a) mean & standard deviation b) range & mode
c) standard deviation & 25th percentile d) range & standard deviation

d) Range and standard deviation (1)

[7] **3.** Identify the following variables as : "*purely categorical (or qualitative)*", "*categorical and ranked*", "*quantitative and discrete*" or "*quantitative and continuous*".

a)	amount of time it takes to commute to work	**quantitative and continuous**
b)	marital status of people	**purely categorical**
c)	number of students in a first grade classroom	**quantitative and discrete**
d)	rating of a professor as: excellent, good, fair, poor	**categorical and ranked**
e)	monthly TV cable bills	**quantitative and continuous**
f)	spring break locations favoured by college students	**purely categorical**
g)	number of cars owned by families	**quantitative and discrete**

[6] **4.** Classify each of the following quantities as either a *parameter* or a *statistic*:

(i) s^2 - **statistics** 1 (ii) β_0 - **parameter** 1 (iii) μ - **parameter** 1

(iv) $\bar{x}$ - **statistics** 1 (v) σ^2 - **parameter** 1 (vi) $\hat{\beta}_0$ - **statistics** 1

[8] **5.** a) Define 2-sided and 1-sided hypotheses and give the steps involved in their testing.

- **<u>2-sided hypothesis</u>: is a 2-tailed test for testing parameter value $\neq 0$ (e.g. $H_0 : \mu = 0$ vs. $H_a : \mu \neq 0$)** 1

- **<u>1-sided hypothesis</u>: is a 1-tailed test for testing parameter value < 0 or > 0** 1/2 1/2
  e.g. $H_0 : \mu \leq 0$ vs. $H_a : \mu > 0$,
  or $H_0 : \mu \geq 0$ vs. $H_a : \mu < 0$

- **<u>steps involved</u>: 1) state H_0 and H_a** 1
  2) test-statistics 1
  3) rejection (critical) region 1
  4) conclusion 1

b) For any hypothesis test, what are the two types of error that may be made?

<u>Type I error</u> = error we make when we reject H_0 when it is true.
P[Type I error] = α 1

<u>Type II error</u> = error we make when we do not reject H_0 when it is false.
P[Type II error] = β 1

[6] **6.** Find the following values from the tables:

a) $z_{0.998}$ b) $z_{0.7291}$ c) $z_{0.002}$ d) $t_{15;0.10}$ e) $-t_{15;0.10}$ f) $t_{15;0.90}$

a) $z_{0.998} = -z_{0.002} = -$ **<u>2.88</u>** 1

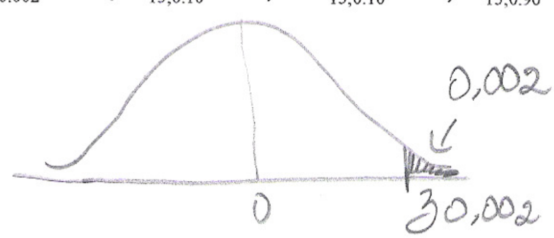
b) $z_{0.7291} = -$ **<u>0.61</u>** 1

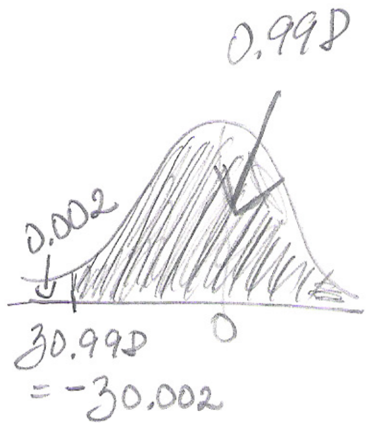
c) $z_{0.002} =$ **<u>2.88</u>** 1

d) $t_{15;0.10} =$ **<u>1.341</u>** 1

e) $-t_{15;0.10} = -$ **<u>1.341</u>** 1

f) $t_{15;0.90} = -t_{15;0.10} = -$ **<u>1.341</u>** 1





[6] 7. Consider a normal population distribution with the value of σ known.

a) What is the confidence level for the interval

(i) $\bar{x} \pm 1.96\sigma/\sqrt{n}$ $\Rightarrow z_{\alpha/2} = \mathbf{1.96} \Rightarrow \alpha/2 = \mathbf{0.025} \Rightarrow \alpha = \mathbf{0.05} \Rightarrow 1-\alpha = \mathbf{0.95}$

$\therefore$ **<u>95% C.I. for</u>** μ (1)

(ii) $\bar{x} \pm 2.58\sigma/\sqrt{n}$ $\Rightarrow z_{\alpha/2} = \mathbf{2.58} \Rightarrow \alpha/2 = \mathbf{0.005} \Rightarrow \alpha = \mathbf{0.010} \Rightarrow 1-\alpha = \mathbf{0.99}$

$\therefore$ **<u>99% C.I. for</u>** μ (1)

(iii) $\bar{x} \pm 3.09\sigma/\sqrt{n}$ $\Rightarrow z_{\alpha/2} = \mathbf{3.09} \Rightarrow \alpha/2 = \mathbf{0.0010} \Rightarrow \alpha = \mathbf{0.0020} \Rightarrow 1-\alpha = \mathbf{0.998}$

$\therefore$ **<u>99.8% C.I. for</u>** μ (1)

b) What value of z in the confidence interval formula

$$\left(\bar{x} - z_{\alpha/2}\,\sigma/\sqrt{n}, \bar{x} + z_{\alpha/2}\,\sigma/\sqrt{n}\right)$$

results in a confidence level of

(i) 89.68% $\Rightarrow 1-\alpha = \mathbf{0.8968} \Rightarrow \alpha = \mathbf{0.1032} \Rightarrow \alpha/2 = \mathbf{0.0516} \Rightarrow z_{\alpha/2} = \underline{\mathbf{1.63}}$ (1)

(ii) 99.20% $\Rightarrow 1-\alpha = \mathbf{0.9920} \Rightarrow \alpha = \mathbf{0.0080} \Rightarrow \alpha/2 = \mathbf{0.0040} \Rightarrow z_{\alpha/2} = \underline{\mathbf{2.65}}$ (1)

(iii) 75.40% $\Rightarrow 1-\alpha = \mathbf{0.7540} \Rightarrow \alpha = \mathbf{0.2460} \Rightarrow \alpha/2 = \mathbf{0.1230} \Rightarrow z_{\alpha/2} = \underline{\mathbf{1.16}}$ (1)

[3] 8. Given that the sample variance is defined by

$$s^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

where $x_1, \ldots, x_n$ is a sample, $\bar{x}$ the sample mean. Show that

$$s^2 = \frac{1}{(n-1)}\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right).$$

Solution:

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 \overset{1}{=} \sum_{i=1}^{n}\left(x_i^{\,2}+\bar{x}^2-2\bar{x}x_i\right)^2 = \sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}\bar{x}^2 - 2\,\bar{x}\sum_{i=1}^{n}x_i =$$

$$\sum_{i=1}^{n}x_i^2+n\bar{x}^2 - 2\,\bar{x}\sum_{i=1}^{n}x_i = \sum_{i=1}^{n}x_i^2+n\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n^2} - 2\,\frac{\sum_{i=1}^{n}x_i}{n}\sum_{i=1}^{n}x_i =$$

$$\sum_{i=1}^{n}x_i^2+\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n} - 2\,\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n} \overset{1}{=} \sum_{i=1}^{n}x_i^2 - \frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}$$

1

i.e. $$\frac{1}{(n-1)}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{(n-1)}\left(\sum_{i=1}^{n}x_i^2-\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}\right)$$ **Q.E.D.**