/

Fall 2011, MT 2 Solutions

MATH 211

Question 1: (20 Points) A machine puts an average of μ milliliters (mL) of dye (boya) per can of paint. The amount of dye discharged is known to have a normal distribution with a standard deviation of 0.4 mL. If more than 6 mL of dye are discharged when making paint, then it is unacceptable.

a) (7 points) Determine μ so that only 1% of the cans of paint will be unacceptable.

$$X \sim N(\mu, 0.4^2)$$

table

$$\mathbb{P}\{X > 6\} = \mathbb{P}\{Z > \frac{6-\mu}{0.4}\} = 0.01 \Rightarrow \frac{6-\mu}{0.4} = 2.33$$

$$6-\mu = 0.932$$

$$\boxed{\mu = 5.068}$$

b) (6 points) What is the probability that a can of paint has at least 4.5 mL of dye?

$$\mathbb{P}\{X \geq 4.5\} = \mathbb{P}\{Z \geq \frac{4.5-5.068}{0.4}\}$$

$$= \mathbb{P}\{Z \geq -1.42\}$$

$$= 1 - \mathbb{P}\{Z \leq -1.42\} = 1 - 0.0778 = \boxed{0.9222}$$

c) (7 points) If a random sample of 16 cans is taken, what is the probability that the sample mean is between 5 and 6?

$$\mathbb{P}\{5 < \bar{X} < 6\} = \mathbb{P}\{\frac{5-5.068}{0.4/\sqrt{16}} < Z < \frac{6-5.068}{0.4/\sqrt{16}}\}$$

$$= \mathbb{P}\{-0.68 < Z < 9.32\}$$

$$\cong 1 - \mathbb{P}\{Z \leq -0.68\} = 1 - 0.2483$$

$$= \boxed{0.7517}$$

Question 2: (20 Points) An appropriate model for income is Pareto distribution, which has the probability density function

$$f_Y(y;\theta) = \theta\left(\frac{1}{y}\right)^{\theta+1} \qquad y \geq 1$$

Suppose we have a data set consisting of a random sample of size n.

a) (10 points) Find the maximum likelihood estimator for θ.

$$L(\theta) = \prod_{i=1}^{n} \theta\left(\frac{1}{y_i}\right)^{\theta+1} = \theta^n \frac{1}{y_1^{\theta+1}} \cdots \frac{1}{y_n^{\theta+1}}$$

$$\ln L(\theta) = n \ln \theta - (\theta+1)(\ln y_1 + \dots + \ln y_n)$$

$$\frac{d \ln L(\theta)}{d\theta} = \frac{n}{\theta} - (\ln y_1 + \dots \ln y_n) = 0$$

for finding $\hat{\theta}$

$$\Rightarrow \boxed{\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \ln Y_i}}$$

b) (10 points) Find the method of moments estimator for θ.

$$\mathbb{E}Y = \theta \int_1^{\infty} y \cdot \left(\frac{1}{y}\right)^{\theta+1} dy = \theta \cdot \frac{y^{-\theta+1}}{-\theta+1}\Big|_1^{\infty}$$

$$= 0 - \frac{\theta}{1-\theta} = \frac{\theta}{\theta-1}$$

$$\frac{\hat{\theta}}{\hat{\theta}-1} = \bar{Y} \Rightarrow \hat{\theta} = \bar{Y}\hat{\theta} - \bar{Y}$$

$$\hat{\theta}(1-\bar{Y}) = -\bar{Y}$$

$$\Rightarrow \boxed{\hat{\theta} = \frac{\bar{Y}}{\bar{Y}-1}}$$

Question 3: (20 Points) During the competition of the brands of beers "Efes Pilsen" and "Tuborg", a taste test has been performed. A random sample of 80 beer drinkers participated in this test and tasted two glasses of beers without knowing what brand was in which. It turned out that 48 of the subjects liked Efes better.

a) (4 points) Let $X_1, \dots, X_{80}$ correspond to the sample described in the question such that $X_i = 1$ if subject i likes Efes better, and $X_i = 0$ otherwise. Write an estimator $\hat{p}$ for the proportion of all beer drinkers who like Efes better, as a function of this sample; then find the estimate p_e using the information given in the question.

$$\hat{p} = \frac{X_1 + \dots + X_{80}}{80}, \quad p_e = \frac{48(1) + 32(0)}{80} = \boxed{0.6}$$

b) (4 points) Is your estimator in part a) unbiased? Show.

$$\mathbb{E}\left(\frac{X_1 + \dots + X_{80}}{80}\right) = \frac{1}{80}\left(\mathbb{E}(X_1) + \dots + \mathbb{E}(X_n)\right)$$

$$= \frac{1}{\cancel{80}}\cancel{80} \cdot \mathbb{E}(X_i) = 1 \cdot p + 0 \cdot (1-p) = \boxed{p}$$

$$\Rightarrow \boxed{\hat{p} \text{ is unbiased.}}$$

c) (7 points) Construct a 90% confidence interval for the proportion of all beer drinkers who like Efes better.

$$p_e \pm z_{\alpha/2}\sqrt{\frac{p_e(1-p_e)}{n}} = 0.6 \pm 1.65\sqrt{\frac{(0.6)(0.4)}{80}}$$

$$\Rightarrow \boxed{[0.51, 0.69]}$$

d) (5 points) In repeated sampling, suppose 20 confidence intervals, all with 90% confidence, are constructed from different samples with varying sample sizes.

- On the average, how many of these intervals do you expect to cover the population proportion?

$$20 \cdot 0.9 = \boxed{18}$$

- Will the width of these intervals be the same? Why or why not?

$$\text{width} = z_{\alpha/2}\sqrt{\frac{p_e(1-p_e)}{n}}$$

n and p_e can change

$$\Rightarrow \boxed{\text{width can change}}$$

Question 4: (20 Points) Lasers are used to inspect the solder joints (lehimli kısımlar) in circuit boards (elektronik devre kartları). The manufacturer of an inspection equipment based on lasers claims that its product can inspect 10 solder joints per second on average. The inspection equipment was tested on 35 different circuit boards by operating the equipment for exactly one second and counting the number of solder joints inspected. The average number of solder joints inspected was found to be 9.8 . Assume that the population standard deviation is 0.7 .

a) (10 points) Do you doubt the manufacturer's claim? Test at $\alpha = 0.01$.

$H_0 : \mu = 10$

$H_1 : \mu \neq 10$

$$z^* = \frac{9.8 - 10}{0.7/\sqrt{35}} = -1.69$$

$\alpha/2 = 0.005$ and $z_{0.005} = 2.58$

$\Rightarrow |z^*| < |z_{0.005}| \Rightarrow$ don't reject H_0

So, $\nexists$ significant evidence to doubt the claim.

b) (4 points) Find the P-value.

$$\text{P-value} = 2 \cdot \mathbb{P}\{z < -1.69\} = 2 \cdot 0.04555 = \boxed{0.091}$$

c) (6 points) Find the Type II error if the true mean number of solder joints that can be inspected is really equal to 9.5

$$\frac{\bar{X} - 10}{0.7/\sqrt{35}} = \pm 2.58 \Rightarrow \bar{x} = 10.3, 9.7$$

$\mathbb{P}\{\text{fail to reject } H_0 \mid \text{true mean is } 9.5\}$

$$= \mathbb{P}\left\{\frac{9.7 - 9.5}{0.7/\sqrt{35}} < z < \frac{10.3 - 9.5}{0.7/\sqrt{35}}\right\}$$

$$= \mathbb{P}\{1.61 < z < 6.85\}$$

$$\cong \mathbb{P}\{1.61 < z\} = \boxed{0.0537}$$

Question 5: (20 Points) Synechocystis, a bacterium, is used by scientists to model DNA behavior. Genes of the bacterium responsible for photosynthesis are isolated and investigated for sensitivity to light. The following are standardized growth measurements of several gene clusters in dark:

1.41 1.83 -0.79 -1.21 1.71 2.14 1.04 -0.21 1.3 1.93

Assume the standardized growth measurements are normally distributed with variance 1.44 .

a) (6 points) Find a 99% confidence interval for the mean standardized growth.

$\bar{X} = 0.915$, $\alpha = 0.01$, $z_{\alpha/2} = 2.58$, $\sigma = \sqrt{1.44} = 1.2$

$n = 10$

$$CI = \left[\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} , \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

$$= \boxed{[-0.063, 1.893]}$$

b) (8 points) It is known that the mean standardized growth in light is 0.1 . Is the mean standardized growth significantly higher in dark?

$H_0 : \mu = 0.1$

$H_1 : \mu > 0.1$

$\bar{X} = 0.915$

$$\mathbb{P}\left\{ Z > \frac{0.915 - 0.1}{1.2/\sqrt{10}} \right\} = \mathbb{P}\{Z > 2.15\}$$

$= 0.0158 \leftarrow$ P-value

P-value $< \alpha = 0.05 \Rightarrow$ reject H_0

$\boxed{\text{Yes, significantly higher in dark.}}$

c) (2 points) What is the proportion of the observations larger than 1.5, in the sample? Note that this is an estimate of the probability that standardized growth is larger than 1.5 in the gene population.

$$\frac{4}{10} = \boxed{0.4}$$

d) (4 points) If one wants to estimate the probability that standardized growth is larger than 1.5 in the gene population within 0.06 of the true value of this probability (say p), with 95% confidence, what should the smallest sample size be?
(Hint: You can use your answer in part c) as an estimate of p , or some other appropriate value).

$\alpha = 0.05 \Rightarrow z_{\alpha/2} = 1.96$

need : $1.96 \sqrt{\frac{(0.4)(0.6)}{n}} \leq 0.06$

$\Rightarrow \boxed{n_{min} = 257}$