

**Midterm 2005-05-09 8:00-10:00 PM**

Closed book, closed note. One two-sided, letter-sized, hand-written formula sheet allowed. A simple calculator is allowed. Work all problems. Print your answer legibly on separate sheets of paper. Express numbers as fractions reduced to lowest terms or as decimals correct to three places (or significant digits). Show your work.

- 1 To evaluate the success of a 1-year experimental program designed to increase the mathematical achievement of underprivileged high school seniors, a random sample of participants in the program will be selected and their math scores will be compared with the previous year's statewide average of 525 for underprivileged seniors. The researchers want to determine whether the experimental program has increased the mean achievement level  $\mu$  over the previous year's statewide average. Suppose a random sample of 100 students is selected yielding the sample mean 542 and  $s=76$ .
  - a) Is there a sufficient evidence to conclude that the mean math achievement level has been increased? Set up the null and alternative hypotheses,  $H_0$  and  $H_a$ . Test it at 5% significance level.
  - b) Compute the P-value for the above hypotheses.
  - c) What're the advantages of using P-value in place of reporting the results as in a)?
  - d) Suppose the true  $\mu$  is 540. What is the power of the test in a) for this particular alternative?
  - e) Other things being equal, what would happen to the power you computed above in d) if the true  $\mu$  is less than 540 (but larger than 525)? Will it increase or decrease? Explain why.
  - f) Other things being equal, what would happen to the power you computed above in d) if the sample size is smaller than 100? Will it increase or decrease? Explain why.
  - g) Construct 95% confidence interval on  $\mu$ .
  - h) You want the 95% confidence interval above to have width less than 30. Other things being equal, what is the smallest sample size you need?
  - i) Suppose the study above was a pilot study. Now, in designing the main study, for the test at level  $\alpha=.05$ , what sample size is needed to have a probability of Type II error of at most .025 if the actual mean is increased to 550?
  - j) Can you still apply above tests and confidence intervals if sample size is small, say less than 30? What further condition(s) do you need to check to apply them and how do you check the condition(s)?
  - k) Suppose, instead of the above, we want to detect the change (either increase or decrease) in the mean math score compared to the previous year's score. How does the P-value in b) change?
  
- 2 For a one-sided statistical hypothesis testing (of the form  $H_0: \mu \leq \mu_0$  vs.  $H_a: \mu > \mu_0$ ) in a certain hypothesis testing, you obtain the P-value of 0.02. Answer the following questions.
  - a) Write down a 1-sentence definition/interpretation of the P-value in statistical hypothesis testing.
  - b) Can you reject this hypothesis at 5% significance level? How about 1% significance level?
  - c) What will be the P-value for the two-sided hypothesis (of the form  $H_0: \mu = \mu_0$  vs.  $H_a: \mu \neq \mu_0$ )?

(Over, please)

- 3 You want to inspect whether a die is fair, in particular, whether the probability of getting six-dots is larger than  $1/6$ . You toss it 100 times to check it out.
- Let  $Y$  be the number of six-dots you get in 100 tosses. If the die was indeed fair, what is the distribution of  $Y$ ? Specify the name of the distribution and values of parameters.
  - Approximate the distribution in a) by a normal distribution, again assuming the die was fair. Specify the mean and standard deviation.
  - At the end of the experiment, you have observed 20 six-dots. If the true probability of six-dots were  $1/6$ , what would be the chance that the observed number of six-dots is greater than or equal to 20? Use the approximate distribution you gave in b) to answer this.
  - The quantity you computed above is the P-value for the following hypotheses:  $H_0: P(\text{observing six-dots}) \leq 1/6$  vs.  $H_a: P(\text{observing six-dots}) > 1/6$ . Test the hypothesis at 5% significance level using the computed P-value.
  - (20 pt; Bonus question, required for 4612 students) Consider a similar experiment but assume you tossed 1,000 times (instead of 100) and observed 200 six-dots. Compute the P-value for the same hypotheses, following similar steps as above. Can you reject the null hypothesis at 1% significance level?

### Midterm Solution

1  $\mu_0=525$  ,  $n=100$ ,  $\bar{y}=542$ ,  $s=76$ .

- a) Is there a sufficient evidence to conclude that the mean math achievement level has been increased?  
Set up the null and alternative hypotheses,  $H_0$  and  $H_a$ . Test it at 5% significance level.

$H_0: \mu \leq 525$  vs.  $H_a: \mu > 525$ .

Reject  $H_0$  when  $z > -z_{.05} = 1.64$ .

The observed  $z$  is  $z = \frac{\bar{y} - \mu_0}{\sigma / \sqrt{n}}$ , approximated by  $(542-525)/7.6 = 2.23$ . Reject  $H_0$ . (There's sufficient

evidence that the new program has increased the mean math score)

- b) Compute the P-value for the above hypotheses.

$P(z > \text{computed } z, \text{ given } H_0 \text{ is true}) = P(z > 2.23) = 1 - \text{pnorm}(2.23)$  (In R function) = 0.0129

- c) What're the advantages of using P-value in place of reporting the results as in a)?

Show the strength of the evidence in the data against the (null) hypothesis rather than a binary result of accept or reject; doesn't need to use a single particular significance level  $\alpha$ .

- d) Suppose the true  $\mu$  is 540. What is the power of the test in a) for this particular alternative?

$$\beta(540) = P\left(Z \leq z_\alpha - \frac{|\mu_0 - \mu_1|}{\sigma_{\bar{y}}}\right) = P(Z \leq 1.64 - \text{abs}(525-540)/7.6) = \text{pnorm}(-0.334) = .369$$

$$\text{PWR}(540) = 1 - \beta(540) = .631$$

- e) Other things being equal, what would happen to the power you computed above in d) if the true  $\mu$  is less than 540 (but larger than 525)? Will it increase or decrease? Explain why.

$$\mu_1 \text{ decreases} \Rightarrow |\mu_0 - \mu_1| \text{ decreases} \Rightarrow z_\alpha - \frac{|\mu_0 - \mu_1|}{\sigma_{\bar{y}}} \text{ increases} \Rightarrow \beta(\mu_1) \text{ increases} \Rightarrow \text{PWR}$$

decreases

Another way to see this is to note that the problem of telling  $H_0$  and an  $H_a$  apart becomes harder (thus power gets smaller) if the true mean  $\mu_1$  in  $H_a$  is close to the null mean, i.e.  $\mu_1$  is small.

- f) Other things being equal, what would happen to the power you computed above in d) if the sample size is smaller than 100? Will it increase or decrease? Explain why.

In a similar logic as above, the power of the test decreases if the sample size decreases (less information or less power in smaller sample)

- g) Construct 95% confidence interval on  $\mu$ .

$$\bar{y} \pm z_{.025} \sigma_{\bar{y}}, \text{ which is approximated by } c(542 - 1.96*7.6, 542 + 1.96*7.6) = (527, 557)$$

- h) You want the 95% confidence interval above to have width less than 30. Other things being equal, what is the smallest sample size you need?

$$\text{Use } E=30/2 \text{ in the formula } n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2} = 1.96^2 * 76^2 / 15^2 = 98.6. \text{ The smallest sample size is 99.}$$

- i) Suppose the study above was a pilot study. Now, in designing the main study, for the test at level  $\alpha=.05$ , what sample size is needed to have a probability of Type II error of at most .025 if the actual mean is increased to 550?

Use the formula  $n = \frac{\sigma^2}{\Delta^2} (z_\alpha + z_\beta)^2 = 76^2 / (550 - 525)^2 * (z_{0.05} + z_{0.025})^2$

$= 76^2 / (550 - 525)^2 * (1.64 + 1.96)^2 = 119.77$

So, we need sample size of at least 120.

- j) Can you still apply above tests and confidence intervals if sample size is small, say less than 30? What further condition(s) do you need to check to apply them and how do you check the condition(s)?

We can apply them if the distribution of the sample is close to normal. Need to use the normal probability plot and check if data points form a straight line.

- k) Suppose, instead of the above, we want to detect the change (either increase or decrease) in the mean math score compared to the previous year's score. How does the P-value in b) change?

The P-value becomes  $2P(z > |\text{computed } z|, \text{ given } H_0 \text{ is true}) = 2P(z > 2.23) = 2 * 0.0129 = 0.0258$

- 2 For a one-sided statistical hypothesis testing (of the form  $H_0: \mu \leq \mu_0$  vs.  $H_a: \mu > \mu_0$ ) in a certain hypothesis testing, you obtain the P-value of 0.02. Answer the following questions.

- a) Write down a 1-sentence definition/interpretation of the P-value in statistical hypothesis testing.

P-value is the probability that the test statistic shows evidence that is as much or more against the null hypothesis as the observed test statistic.

- b) Can you reject this hypothesis at 5% significance level? How about 1% significance level?

Cannot reject at 5%. Can reject at 1%. (P-value is the smallest significance level at which you can REJECT the null)

- c) What will be the P-value for the two-sided hypothesis (of the form  $H_0: \mu = \mu_0$  vs.  $H_a: \mu \neq \mu_0$ )?

In many situations, two-sided P-values are twice the one-sided P-value. The answer is  $2 * 0.02 = 0.04$ . You can also observe this in problem 1- k) above.

- 3 You want to inspect whether a die is fair, in particular, whether the probability of getting six-dots is larger than  $1/6$ . You toss it 100 times to check it out.

- a) Let Y be the number of six-dots you get in 100 tosses. If the die was indeed fair, what is the distribution of Y? Specify the name of the distribution and values of parameters.

$n=100, p=1/6. Y \sim \text{bin}(100, 1/6)$

- b) Approximate the distribution in a) by a normal distribution, again assuming the die was fair. Specify the mean and standard deviation.

The mean is  $np=100/6 = 16.7$  and  $SD=\sqrt{100*1/6*5/6} = 3.73. Y \sim N(16.7, 3.73^2)$ .

- c) At the end of the experiment, you have observed 20 six-dots. If the true probability of six-dots were  $1/6$ , what would be the chance that the observed number of six-dots is greater than or equal to 20? Use the approximate distribution you gave in b) to answer this.

The Z-score is  $z=(20-16.7)/3.73 = 0.885$ . Thus  $P(Z > 0.885) = 1 - \text{pnorm}(0.885) = 0.188$

- d) The quantity you computed above is the P-value for the following hypotheses:  $H_0: P(\text{observing six-dots}) \leq 1/6$  vs.  $H_a: P(\text{observing six-dots}) > 1/6$ . Test the hypothesis at 5% significance level using the computed P-value.

Cannot reject the null hypothesis since P-value  $> .05$ .

- e) (20 pt; Bonus question, required for 4612 students) Consider a similar experiment but assume you tossed 1,000 times (instead of 100) and observed 200 six-dots. Compute the P-value for the same hypotheses, following similar steps as above. Can you reject the null hypothesis at 1% significance level?

The distribution is  $Y \sim \text{bin}(1000, 1/6)$ .

The mean is  $np = 1000/6 = 167$  and  $SD = \sqrt{1000 * 1/6 * 5/6} = 11.8$ .  $Y \sim N(167, 11.8^2)$ .

The Z-score is  $z = (200 - 167) / 11.8 = 2.80$ . Thus  $P(Z > 2.80) = 1 - \text{pnorm}(2.80) = 0.00256$

Can reject the null hypothesis since P-value  $< .01$ .