Math 171.04 Spring 2004
April 9, 2004
Topics Covered for Prelim #2

Chapter 3: The Normal Distribution

-If X has a normal distribution with mean μ and standard deviation σ (that is, variance σ^2) we write $X \sim \mathcal{N}(\mu, \sigma)$.

-If $X \sim \mathcal{N}(\mu, \sigma)$ then

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

This allows us to use Table A to look up normal probabilities *for any* normal random variables.

-If $Z \sim \mathcal{N}(0, 1)$ then

$$X = \sigma Z + \mu \sim \mathcal{N}(\mu, \sigma).$$

-If $X_1, X_2, \ldots, X_n$ are independent $\mathcal{N}(\mu, \sigma)$ random variables, then

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \sim \mathcal{N}(\mu, \frac{\sigma}{\sqrt{n}}).$$

-If $X_1 \sim \mathcal{N}(\mu_1, \sigma_1)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2)$ are independent, then

$$E(X_1 + X_2) = E(X_1) + E(X_2) = \mu_1 + \mu_2$$

and

$$\mathrm{SD}(X_1 + X_2) = \sqrt{\mathrm{Var}(X_1 + X_2)} = \sqrt{\mathrm{Var}(X_1) + \mathrm{Var}(X_2)} = \sqrt{\sigma_1^2 + \sigma_2^2}.$$

Furthermore, the *distribution* of $X_1 + X_2$ is normal; that is,

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}).$$

Exercise: Find the distribution of $X_1 - X_2$. (Hint: It is normal.)

Chapters 4 and 5: Linear Regression and Correlation

-explanatory (independent) and response (dependent) variables

-The coefficient of correlation is

$$r = \frac{1}{n-1} \sum \left(\frac{x_i - \overline{x}}{s_x} \right) \left(\frac{y_i - \overline{y}}{s_y} \right) = \frac{1}{(n-1)\, s_x\, s_y} \sum \left(x_i - \overline{x} \right) \left(y_i - \overline{y} \right).$$

-The equation of the regression line is

$$\hat{y} = a + bx$$

where

$$b = r \frac{s_y}{s_x} \quad \text{and} \quad a = \overline{y} - b\overline{x}.$$

Chapters 7 and 8: Producing Data

-population, sample, simple random sample (i.i.d. random variables), voluntary response sample, stratified random sample

-observational study vs. experimental study

-confounding, bias, undercoverage, nonresponse

-random digits, "label and table"

-subjects and treatments

-Use Table B to randomly assign subjects to control group (placebo) or treatment group.

-comparative experiments, randomized comparative experiments, completely randomized design

-double blind

-matched pairs design, block design

Chapter 13: Confidence Intervals

-A level C confidence interval for μ if the standard deviation σ is known is

$$\overline{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

where z^* is the critical value corresponding to C and is found from Table A.

-Notice that a confidence interval for μ is composed of a *point estimate* (namely $\overline{x}$) and a *margin of error.*

-If we want a desired margin of error m, then we need at least n data points, where n satisfies

$$m = z^* \frac{\sigma}{\sqrt{n}}.$$

That is,

$$n = \left(z^* \frac{\sigma}{m} \right)^2.$$

Note that n *must* be a whole number.

Chapter 14: Hypothesis Testing

-Null hyposthesis: $H_0 : \mu = \mu_0$ for some fixed μ_0

-(a) Alternative hypothesis: $H_a : \mu \neq \mu_0$ (two-sided)

-(b) Alternative hypothesis: $H_a : \mu > \mu_0$ (one-sided)

-(c) Alternative hypothesis: $H_a : \mu < \mu_0$ (one-sided)

-For a given significance level α, we can determine when we reject H_0.

-Assume H_0 is true. Compute the test statistic

$$z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}.$$

-compute the P-value:
(a) $P = 2P(|Z| \geq z)$
(b) $P = P(Z \geq z)$
(c) $P = P(Z \leq z)$
where $Z \sim \mathcal{N}(0, 1)$ so that Table A can be used.

EQUIVALENTLY, if you are given a rejection rule in terms of $\overline{x}$, then you can find the significance level α from Table A.

Chapter 15: Inference in Practice

-look at figure 15.2 on page 378. A Type I error is made when we reject H_0 but H_0 is true. A Type II error is made when we accept H_0 but H_a is true (i.e., when we do not reject H_0 when H_a is true).

-If (ℓ, r) is a level C confidence interval for μ, then there is a natural hypothesis test associated. The test of $H_0 : \mu = \mu_0$ vs. $H_a : \mu \neq \mu_0$ has significance level $\alpha = 1 - C$ and rejects H_0 if μ_0 does NOT lie in the interval (ℓ, r).

$$\alpha = P(\text{Type I error}) = P(\text{reject } H_0 | H_0 \text{ is true}) = P(\text{reject } H_0 | \mu = \mu_0)$$

-The probability that a fixed level α test will reject H_0 when a particular alternative value of the parameter is true is called the power of the test against that alternative.

$$\text{Power} = 1 - P(\text{Type II error for that alternative})$$

-Assume we are testing $H_0 : \mu = \mu_0$ vs. $H_a : \mu > \mu_0$. We find our rejection region: reject H_0 if $\overline{x} > a$. Suppose that ONE plausible alternative is $H_a : \mu = \mu_1$. The probability of a Type II error for this alternative is

$$P(\text{fail to reject } H_0 | \mu = \mu_1) = P(\overline{x} \leq a | \mu = \mu_1)$$

which can now be computed from Table A.

Practice Problems

Old Prelim Number 1 and Old Prelim Number 2 both provide good practice with this material. However, there is not enough on Type I/II errors and power on them.

Exercise 14.35, page 360
Exercise 15.10, page 377
Exercise 15.11, page 377
Exercise 15.12, page 377
Exercise 15.14, page 380
Exercise 15.15, page 380
Read Example 14.7, page 351
Exercise 15.32, page 385