Topic 3: Basic concepts – part 2

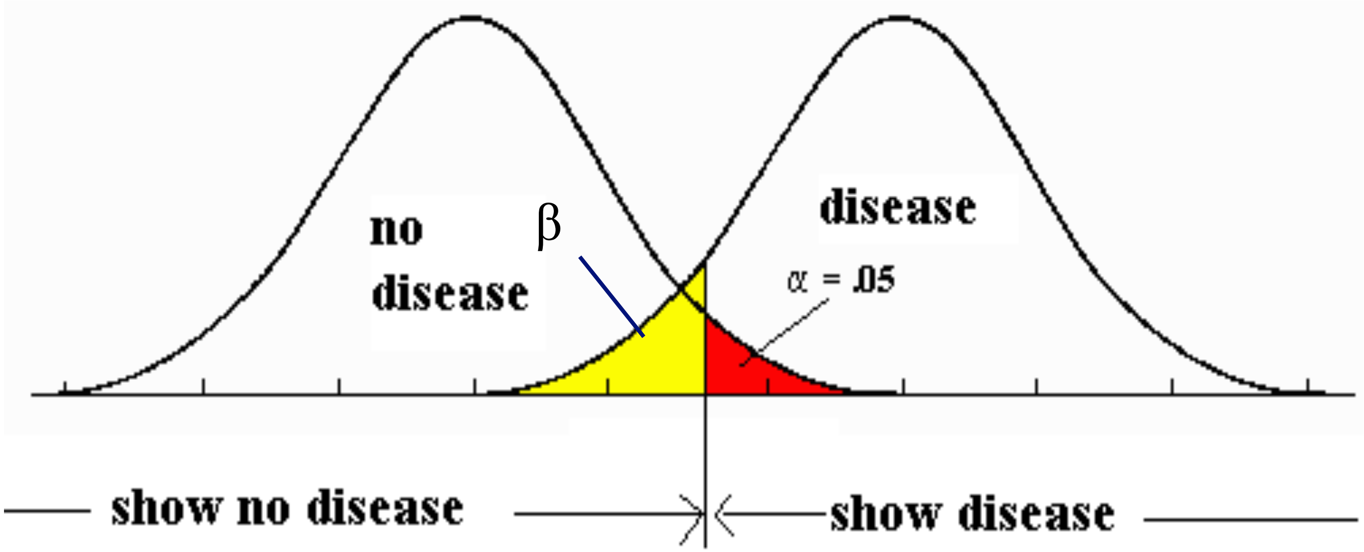
Montgomery: chapter 2

Prof. Lingling An
University of Arizona

Outline

- Type I, II error
- Choice of sample size
- Summary of tests on means & variances
- Confidence interval and normality check

Illustration of Types of error



Two types of error

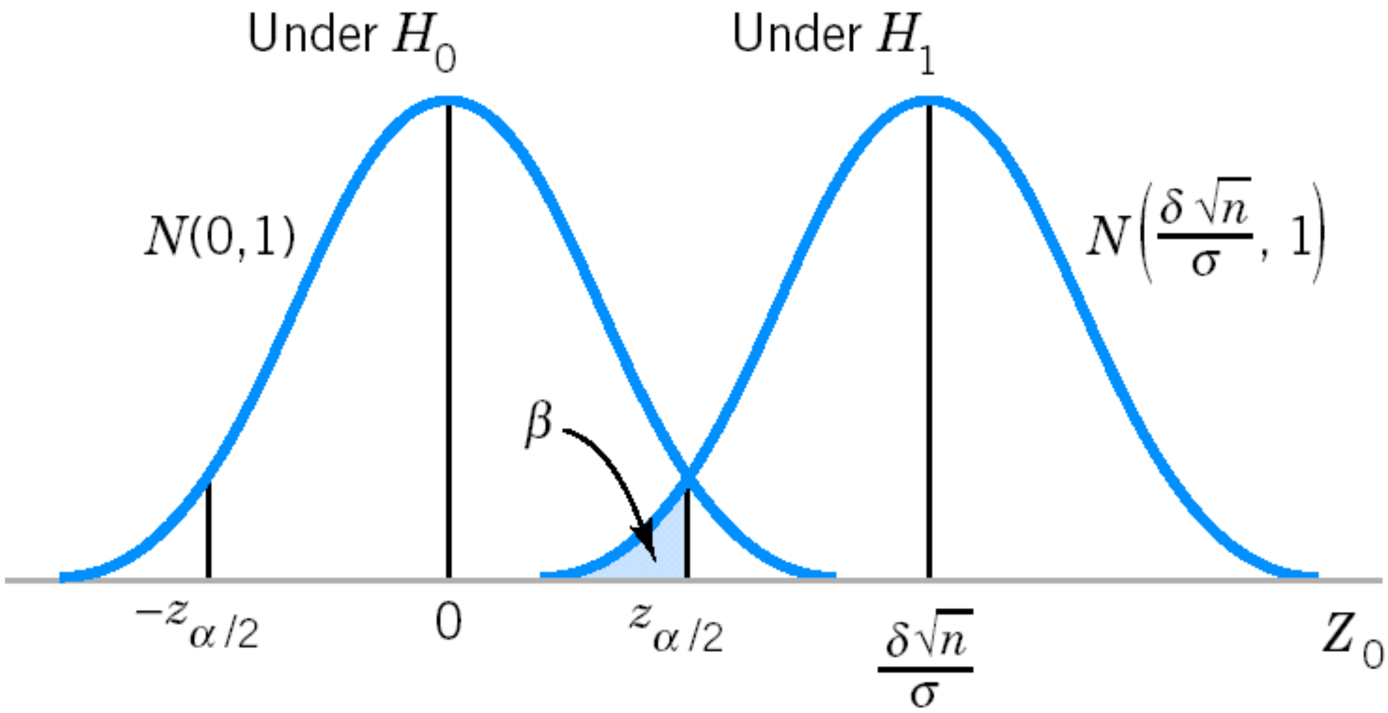
Truth / Data	H_0 Correct (no disease)	H_1 Correct (disease)
Decide H_0 "fail to reject H_0" (test shows no disease)	1- α True Negative	β False Negative
Decide H_1 "reject H_0" (test shows disease)	α False Positive	1- β True Positive

- Type I error - False positive rate
- α = P(reject H_0 | H_0 true)
  - Probability reject the true null hypothesis
- α is significance level

- Type II error--False negative rate
- β = P(do not reject H_0 | H_1 true)
  - Probability not reject a false null hypothesis
- **Power** = 1-β = P(reject H_0 | H_1 true)

One sample test – type I, II error

- H_0: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$, and assume variance σ^2 is known
- Let $\delta = \mu - \mu_0$
- Let $Z_0 = (\bar{x} - \mu_0)/(\sigma/\sqrt{n})$, then under H_0: $Z_0 \sim N(0,1)$ and under H_1: $Z_0 \sim N(\delta\sqrt{n}/\sigma, 1)$



- Type II error depend on:
  - *α*
  - sample size
  - population variance
  - difference between actual and hypothesized means
- To decrease β (type II error) for a fixed α (type I error), we need increase sample size n.

- Why we have such distributions in the previous slide?

Consider the two-sided hypothesis

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

Suppose that the null hypothesis is false and that the true value of the mean is $\mu = \mu_0 + \delta$, say, where $\delta > 0$. The test statistic Z_0 is

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\overline{X} - (\mu_0 + \delta)}{\sigma/\sqrt{n}} + \frac{\delta\sqrt{n}}{\sigma}$$

Therefore, the distribution of Z_0 when H_1 is true is

$$Z_0 \sim N\left(\frac{\delta\sqrt{n}}{\sigma}, 1\right)$$

Choice of sample size (one sample)

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right)$$

$$\boxed{n \simeq \frac{(z_{\alpha/2} + z_{\beta})^2 \sigma^2}{\delta^2}}$$

Choice for sample size (two samples case, known variances)

- H_0: $\mu_1 = \mu_2 + \Delta_0$, H_1: $\mu_1 \neq \mu_2 + \Delta_0$
- Assume two samples have the same size n, then

$$n \simeq \frac{(z_{\alpha/2} + z_{\beta})^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2}$$

  where Δ is the true difference in means.
- An alternative way for calculating sample size is **operating characteristic curve** (read p41-42)

- For *one-sided test* of two-sample case,

$$n = \frac{(z_\alpha + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2}$$

Example

- A product developer is interested in reducing the drying time of a primer paint. Two formulations of the paint are tested: formulation 1 is the standard chemistry, and formulation2 has a new drying ingredient that should reduce drying time.
- From experience, it is known that the standard deviation of drying time is 8 minutes, and this inherent variability should be unaffected by the addition of the new ingredient.
- Suppose that if the true difference in drying times is as much as 10 minutes. He wants to detect this with probability at least 0.90.

- Under the null hypothesis Δ_0 =0 and one-sided alternative hypothesis is with Δ=10. Since the power is 0.9, β=0.10 ($Z_\beta = Z_{0.10}$=1.28). Given α=0.05 (so Z_α=1.645),

$$n = \frac{(Z_\alpha + Z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2} = \frac{(1.645 + 1.28)^2(8^2 + 8^2)}{(10-0)^2} = 11$$

Summary of tests

■ TABLE 2.3

Tests on Means with Variance Known

	Hypothesis	Test Statistic	Fixed Significance Level Criteria for Rejection	*P*-Value
One sample	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$Z_0 = \dfrac{\bar{y} - \mu_0}{\sigma/\sqrt{n}}$	$\lvert Z_0 \rvert > Z_{\alpha/2}$	$P = 2[1 - \Phi(\lvert Z_0 \rvert)]$
	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$		$Z_0 < -Z_\alpha$	$P = \Phi(Z_0)$
	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$		$Z_0 > Z_\alpha$	$P = 1 - \Phi(Z_0)$
two samples	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$Z_0 = \dfrac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$\lvert Z_0 \rvert > Z_{\alpha/2}$	$P = 2[1 - \Phi(\lvert Z_0 \rvert)]$
	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$		$Z_0 < -Z_\alpha$	$P = \Phi(Z_0)$
	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$		$Z_0 > Z_\alpha$	$P = 1 - \Phi(Z_0)$

■ TABLE 2.4

Tests on Means of Normal Distributions, Variance Unknown

One sample

Hypothesis	Test Statistic	Fixed Significance Level Criteria for Rejection	*P*-Value
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$		$\lvert t_0 \rvert > t_{\alpha/2,n-1}$	sum of the probability above t_0 and below $-t_0$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$t_0 = \dfrac{\bar{y} - \mu_0}{S/\sqrt{n}}$	$t_0 < -t_{\alpha,n-1}$	probability below t_0
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$		$t_0 > t_{\alpha,n-1}$	probability above t_0

two samples

Hypothesis	Test Statistic	Fixed Significance Level Criteria for Rejection	*P*-Value
	if $\sigma_1^2 = \sigma_2^2$		
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$t_0 = \dfrac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $v = n_1 + n_2 - 2$	$\lvert t_0 \rvert > t_{\alpha/2,v}$	sum of the probability above t_0 and below $-t_0$
	if $\sigma_1^2 \neq \sigma_2^2$		
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$t_0 = \dfrac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$	$t_0 < -t_{\alpha,v}$	probability below t_0
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$v = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{(S_1^2/n_1)^2}{n_1 - 1} + \dfrac{(S_2^2/n_2)^2}{n_2 - 1}}$	$t_0 > t_{\alpha,v}$	probability above t_0

Summary of tests -2

■ TABLE 2.7

Tests on Variances of Normal Distributions

	Hypothesis	Test Statistic	Fixed Significance Level Criteria for Rejection
One sample	$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$	$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2}$	$\chi_0^2 > \chi^2_{\alpha/2,n-1}$ or $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$
	$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$		$\chi_0^2 < \chi^2_{1-\alpha,n-1}$
	$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$		$\chi_0^2 > \chi^2_{\alpha,n-1}$
two samples	$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$	$F_0 = \frac{S_1^2}{S_2^2}$	$F_0 > F_{\alpha/2,n_1-1,n_2-1}$ or $F_0 < F_{1-\alpha/2,n_1-1,n_2-1}$
	$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 < \sigma_2^2$	$F_0 = \frac{S_2^2}{S_1^2}$	$F_0 > F_{\alpha,n_2-1,n_1-1}$
	$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$	$F_0 = \frac{S_1^2}{S_2^2}$	$F_0 > F_{\alpha,n_1-1,n_2-1}$

Confidence Intervals (See p 44)

- Hypothesis testing gives an objective statement concerning the difference in means, but it doesn't specify "how different" they are.
- **General form** of a confidence interval

$$L \le \theta \le U \text{ where } P(L \le \theta \le U) = 1 - \alpha$$

- The 100(1- α)% **confidence interval** on the difference in two means:

$$\bar{y}_1 - \bar{y}_2 - t_{\alpha/2, n_1+n_2-2} S_p \sqrt{(1/n_1) + (1/n_2)} \le \mu_1 - \mu_2 \le$$

$$\bar{y}_1 - \bar{y}_2 + t_{\alpha/2, n_1+n_2-2} S_p \sqrt{(1/n_1) + (1/n_2)}$$

Example

The actual 95 percent confidence interval estimate for the difference in mean tension bond strength for the formulations of Portland cement mortar is found by substituting in Equation 2.30 as follows:

$$
\begin{aligned}
16.76 - 17.04 - (2.101)0.284\sqrt{\tfrac{1}{10} + \tfrac{1}{10}} \le\ & \mu_1 - \mu_2 \\
\le\ & 16.76 - 17.04 + (2.101)0.284\sqrt{\tfrac{1}{10} + \tfrac{1}{10}}
\end{aligned}
$$

$$-0.28 - 0.27 \le \mu_1 - \mu_2 \le -0.28 + 0.27$$

$$-0.55 \le \mu_1 - \mu_2 \le -0.01$$

Thus, the 95 percent confidence interval estimate on the difference in means extends from -0.55 to -0.01 kgf/cm^2. Put another way, the confidence interval is $\mu_1 - \mu_2 = -0.28 \pm 0.27$ kgf/cm^2, or the difference in mean strengths is -0.28 kgf/cm^2, and the accuracy of this estimate is ± 0.27 kgf/cm^2. Note that because $\mu_1 - \mu_2 = 0$ is *not* included in this interval, the data do not support the hypothesis that $\mu_1 = \mu_2$ at the 5 percent level of significance (recall that the *P*-value for the two-sample *t*-test was 0.042, just slightly less than 0.05). It is likely that the mean strength of the unmodified formulation exceeds the mean strength of the modified formulation.

Checking normal assumptions

- There are two ways of testing normality.
  - Graphical methods:
    - visualize the distributions of random variables or differences between an empirical distribution and a theoretical distribution (e.g., the standard normal distribution).
  - Numerical methods:
    - present summary statistics such as skewness and kurtosis, or conduct statistical tests of normality.

	Graphical Methods	Numerical Methods
Descriptive	Stem-and-leaf plot, (skeletal) box plot, dot plot, histogram	Skewness Kurtosis
Theory-driven	P-P plot Q-Q plot	Shapiro-Wilk, Shapiro- Francia test Kolmogorov-Smirnov test (Lillefors test) Anderson-Darling/Cramer-von Mises tests Jarque-Bera test, Skewness-Kurtosis test

Graphical methods are intuitive and easy to interpret, while numerical methods provide objective ways of examining normality.

Checking normal assumptions – The Normal Probability Plot

$Y_1, Y_2, \ldots, Y_n$ is a random sample from a population with mean μ and variance σ^2.

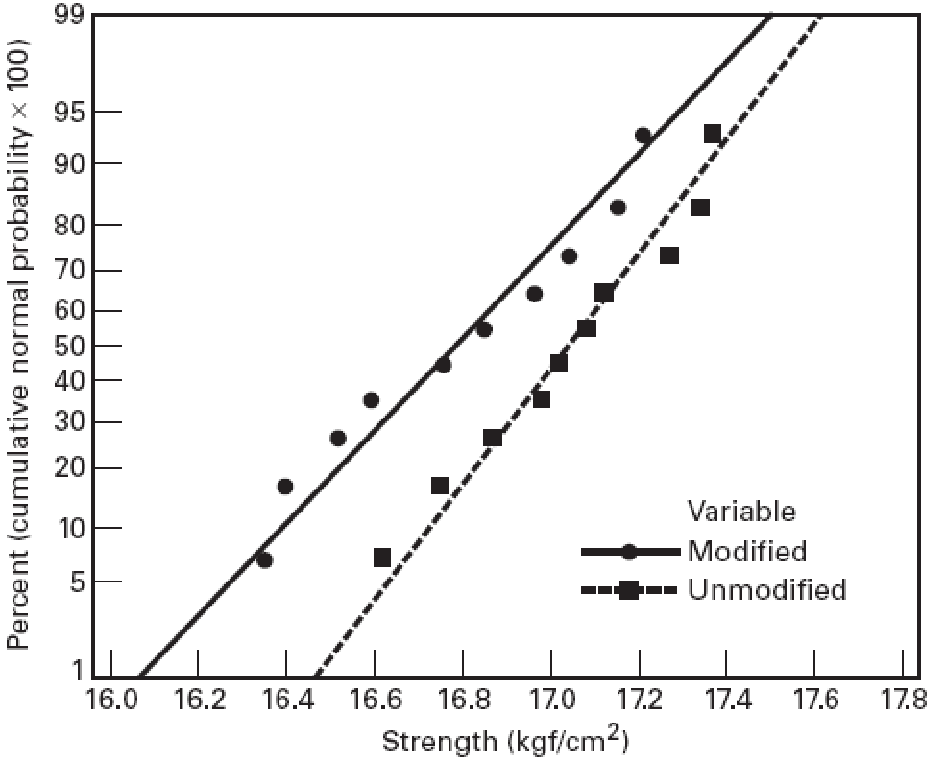
Order Statistics: $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ where $Y_{(i)}$ is the ith smallest value.

if the population is normal, i.e., $N(\mu, \sigma^2)$, then

$E(Y_{(i)}) \approx \mu + \sigma r_{\alpha_i}$ with $\alpha_i = \frac{i-3/8}{n+1/4}$
where r_{α_i} is the 100α_i th percentile of $N(0,1)$ for $1 \leq i \leq n$.

Given a sample $y_1, y_2, \ldots, y_n$, the plot of $(r_{\alpha_i}, y_{(i)})$ is called the normal probability plot

Checking normal assumptions – The Normal Probability Plot



■ FIGURE 2.11 Normal probability plots of tension bond strength in the Portland cement experiment

- the points falling around a straight line indicate normality of the population;
- Deviation from a straight line pattern indicates non-normality

- We discussed method for comparing two conditions or treatments.
- But how about compare more than two conditions or levels of a factor?

Last slide

- Read Sections: finish Ch2