

A Brief Introduction to Intersection-Union Tests

Jimmy Akira Doi
jadoi@unity.ncsu.edu

North Carolina State University
Department of Statistics

Slide 1

Introduction

Often, the quality of a product is determined by several parameters. The product is determined to be acceptable if each of the parameters meets certain standards.

Acceptance Sampling: Upholstery Fabric Example
ASTM Standards (American Society for Testing and Materials)

- Mean breaking strength (in pounds)
- Probability of passing flammability test
- Fabric deemed acceptable when each of these parameters satisfy certain criteria

Slide 2

Hypothesis Testing Framework

- Formulate the null hypothesis H_0 which states that one or more of the parameters **do not meet** their criteria.
- Formulate the alternative hypothesis H_A which states that **all** of the parameters **do meet** their criteria.

Under this framework, note that rejection of the null hypothesis H_0 corresponds to the decision that *the product is acceptable*. Also, the probability of a Type I error will be the *consumer's risk*.

Slide 3

Some Concerns

- Multiple Testing Procedure
Is there a need for a multiplicity adjustment?
- Need to Control Consumer's Risk

Slide 4

Outline

- Review of the Problem of Multiple Comparisons
- Definition of the Intersection-Union Test
- 2 Relevant Theorems
- Example with Simulation Results
- Bioequivalence Introduction

Slide 5

Multiple Comparisons

Example: Experiment with 5 treatments

- 10 possible pairwise comparisons or *contrasts* (*0.05 level t-tests: t_1, t_2, \dots, t_{10}*)
- Combine all tests into ϕ_t , where ϕ_t rejects if any of the t_i rejects.
- Familywise error rate $\alpha^* \neq 0.05$

	<i>n</i>				
	3	5	10	20	45
α^*	14%	23%	40%	64%	90%

- Need multiplicity adjustment

Slide 6

“According to one survey, multiple comparison methods constitute the second most frequently applied group of statistical methods, second only to the F-test ... If they rank second to frequency of use, they rank perhaps first in frequency of abuse.”

Jason Hsu,
Multiple Comparisons: Theory and Methods

Slide 7

Intersection-Union Tests (IUT)

Given $\mathbf{X} \sim f(\mathbf{x}|\theta)$, suppose H_0 is expressed as a union of k sets (*the index set need not be finite*):

$$H_0 : \theta \in \Theta_0 = \bigcup_{i=1}^k \Theta_i \text{ vs. } H_A : \theta \in \Theta_0^c = \bigcap_{i=1}^k \Theta_i^c \quad (1)$$

Suppose for each i , R_i is the rejection region for a test of

$$H_{0i} : \theta \in \Theta_i \text{ vs. } H_{Ai} : \theta \in \Theta_i^c.$$

The rejection region for the IUT of H_0 vs. H_A is $\bigcap_{i=1}^k R_i$.

In other words, the IUT rejects only if all of the tests reject.

Do we need a multiplicity adjustment?

Slide 8

Theorem 1

The following theorems are from: Berger, RL (1997), "Likelihood Ratio Tests and Intersection-Union Tests"

Theorem 1: If R_i is a level- α test of H_{0i} , for $i = 1, \dots, k$, then the IUT with rejection region $R = \cap_{i=1}^k R_i$ is a level- α test of H_0 versus H_A in (1).

(No adjustment necessary!)

Usefulness of IUT: Rely upon the simpler tests of the individual hypotheses and not worry about the joint multivariate dist'n.

Although the IUT above is level- α , the test can be very conservative. Can the IUT also be size- α ?

Slide 9

Proof of Theorem 1

Theorem 1: If R_i is a level- α test of H_{0i} , for $i = 1, \dots, k$, then the IUT with rejection region $R = \cap_{i=1}^k R_i$ is a level- α test of H_0 versus H_A in (1).

Proof: Let $\theta \in \Theta_0 = \cup_{i=1}^k \Theta_i$. So, for some $i = 1, \dots, k$ (say $i = i'$), $\theta \in \Theta_{i'}$. Thus,

$$P_\theta(\cap R_i) \leq P_\theta(R_{i'}) \leq \alpha$$

Since $\theta \in \Theta_0$ was arbitrarily chosen, the IUT is level- α as

$$\sup_{\theta \in \Theta_0} P_\theta(\cap R_i) \leq \alpha$$

(No adjustment necessary)

Slide 10

Theorem 2

Theorem 2: For some $i = 1, \dots, k$, suppose R_i is a size- α rejection region for testing H_{0i} vs. H_{Ai} . For every $j = 1, \dots, k, j \neq i$, suppose R_j is a level- α rejection region for testing H_{0j} vs. H_{Aj} . Suppose there exists a sequence of parameter points $\theta_l, l = 1, 2, \dots$, in Θ_i such that

$$\lim_{l \rightarrow \infty} P_{\theta_l}(R_i) = \alpha,$$

and, for every $j = 1, \dots, k, j \neq i$,

$$\lim_{l \rightarrow \infty} P_{\theta_l}(R_j) = 1.$$

Then the IUT with rejection region $R = \cap_{i=1}^k R_i$ is a size α test of H_0 vs. H_A .

Slide 11

Proof of Theorem 2

Proof:

$$\lim_{l \rightarrow \infty} P_{\theta_l} \left(\bigcap_{m=1}^k R_m \right) \geq \overbrace{\lim_{l \rightarrow \infty} \sum_{m=1}^k P_{\theta_l}(R_m)}^{\text{Bonferroni Inequality}} - (k-1)$$

$$= \alpha + (k-1) - (k-1) = \alpha.$$

Slide 12

Acceptance Sampling: Upholstery Fabric Example

- θ_1 , mean breaking strength (in pounds)
- θ_2 , probability of passing flammability test
- Fabric is acceptable when meeting the standards:
 $\theta_1 > 50$ and $\theta_2 > 0.95$
- $H_0: \{\theta_1 \leq 50 \text{ or } \theta_2 \leq 0.95\}$
 $H_A: \{\theta_1 > 50 \text{ and } \theta_2 > 0.95\}$

X_1, \dots, X_n iid $N(\theta_1, \sigma^2)$, use LRT of $H_{01} : \theta_1 \leq 50$.
 Y_1, \dots, Y_m iid $\text{Bern}(\theta_2)$, use LRT of $H_{02} : \theta_2 \leq 0.95$.

IUT rejection region is $\{(\mathbf{x}, \mathbf{y}) : \frac{\bar{x}-50}{s/\sqrt{n}} > t \text{ and } \sum_{i=1}^m y_i > b\}$.

MC Simulation: $n = m = 58, t = 1.672, b = 57$
 \Rightarrow LRTs are approx. size- α (0.05) tests.

Slide 13

Table 1. Monte Carlo Estimates of α

Fixed θ_2 and Increasing θ_1

	$\theta_1 = 50$	$\theta_1 = 60$	$\theta_1 = 75$	$\theta_1 = 100$
$\theta_2 = .95$.002	.041	.049	.051

Table 2. Monte Carlo Estimates of α

Fixed θ_1 and Increasing θ_2

	$\theta_2 = .95$	$\theta_2 = .99$	$\theta_2 = .999$	$\theta_2 = .9999$
$\theta_1 = 50$.002	.015	.049	.049

Slide 14

Power Analysis

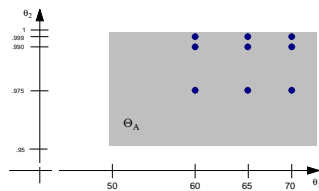


Table 3. MC Estimates of Power

		θ_2		
		.975	.99	.999
θ_1	60	.185	.553	.757
	65	.226	.544	.927
	70	.230	.553	.944

Slide 15

Uses of the *IUT*

- Acceptance Sampling
- Comparisons of Regression Lines
- Testing for Contingency Tables
- Bioequivalence

Slide 16

New Drug Development and Approval

Typically, the approval of a new drug product requires a process that is very time consuming and expensive.

- Phase I, II, III Clinical Trials
- Efficacy, Safety, Risk (side effects)

In the US, if a biopharmaceutical company successfully introduces a new drug into the market, the process typically requires 15 years and approximately \$500 million.

Companies that produce *generic* drugs can bypass this process by showing the drug regulatory agency that their product is *similar* (therapeutically) to that of an approved reference drug. *That is, they want to establish bioequivalence.*

Slide 17

A Definition

(Bio)equivalence

“Two different drugs or formulations of the same drug are called *bioequivalent* if they are absorbed into the blood and become available at the drug action site at about the same rate and concentration.”

Berger and Hsu (1996)

In order for the generic drug to be approved in the US, the Food and Drug Administration (FDA, www.fda.gov) requires a test of the following:

Slide 18

Hypothesis Testing Framework for Bioequivalence (FDA)

Let

μ_R = Mean blood concentration of a reference drug

μ_T = Mean blood concentration of a generic drug

$$H_0 : \mu_T - \mu_R \geq \delta \quad \text{or} \quad \mu_T - \mu_R \leq -\delta$$

$$H_A : \quad -\delta < \mu_T - \mu_R < \delta$$

where $\delta > 0$ is a tolerance limit specified by the drug regulatory agency.

IUT: Combine two, one-sided, size- α tests to obtain an overall size- α test. This test is often referred to as the *TOST*.

Slide 19

TOST

- Most widely used *IUT*
- Much wider recognition of *IUTs* due to this application
- But, *TOST* suffers from a lack of *power*

Slide 20

Summary

- *IUT* avoids the need for multiplicity adjustments.
- 2 relevant theorems:
 - Theorem 1: IUT is level- α when all others are level- α .*
 - Theorem 2: Under certain conditions, IUT is size α .*
- No need to postulate a multivariate model – just combine the simpler individual tests.
- Strongest application found through bioequivalence studies.