STA 106: Applied Statistical Methods: Analysis of Variance

Course Material Summary

University of California at Davis

Last Edit Date: 02/19/2022

STA 106 Summary

Topic	Content
Statistical Concepts	A **population** is the set of all subjects or elements about which we care interested in making decisions (inferences). A **data frame** is a list of the members of the population. A **sample** is a subset of the population which is used to gain insight about the population. Samples are used to represent the larger population of interest. A **parameter** is a numerical descriptive measure of a population. It is usually unknown as the population is too large to be collected. A population can have many parameters, and the parameters can be inferred (or estimated) from a sample (mostly random sample).
Distributions	**Two types of distribution functions:** 1. **Probability density function** (for the continuous random variable) or **probability mass function** (for the discrete random variables). We denote them by pdf or pmf. They describe the probability (or likelihood) of the random variable at a particular value. 2. **Cumulative probability function** (cdf) which is the integral of the probability density function (pdf) or the cumulative sum of the probability mass function (pmf) on a given interval. You have computed p-values for testing the hypothesis, and the p-values are the application of **cumulative probability function**. **Bernoulli distribution:** Bernoulli distribution is the discrete probability distribution of a random variable X which takes the value 1 with probability p and the value 0 with probability 1 − p. For example, toss a coin once. Let X = 1 denote the head and X = 0 the tail. The probability mass function (pmf) of Bernoulli distribution is: $f_X(x) = p^x(1-p)^{1-x},\ x \in \{0,1\}\ and\ p = 0.5$ for tossing a coin once. We can simply denote $X \sim Bernoulli(p)$ $E(X) = \sum x \cdot f_X(x)$ for discrete random variables $E(X) = \int x f_X(x) dx$ for continuous random variables **Binomial distribution:** Binomial distribution describes the probability of getting exactly X = k successes (with probability p for each trial) in n independent Bernoulli trials. For example, we toss the coin 10 times and get 4 heads. The probability mass function (pmf) of Binomial distribution follows: $f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, k \in \{0,1,2,\dots,n\}$ $\#\binom{n}{k} = \frac{n!}{k!(n-k)!}$ The cumulative distribution function can be expressed as:

	$F(k;n,p) = \Pr(X \le k) = \sum_{i=0}^{k} \binom{n}{k} p^k (1-p)^{n-k}, k \in \{0,1,2,\dots,n\}$ We can denote $X \sim Binomial(n,p)$ **Normal distribution:** Normal (or Gaussian) distributions are a type of continuous probability distribution for a real-valued random variable. Normal distributions play a central role in statistical inference! The pdf for a normal random variable X is: $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), x \in (-\infty, \infty)$ The cdf for a normal random variable X is: $F_X(x) = P(X \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right) dz$ We denote $X \sim N(\mu, \sigma^2)$ **Z-score:** $Z = \frac{X-\mu}{\sigma}$ **Central Limit Theorem:** Central Limit Theorem (CLT) Under some conditions, when the sample size n is large ($n \to \infty$), the sample statistic converges to a population parameter (This is a "dirty" definition of CLT). t-distribution: 1. Degree of freedom (df) = the maximum number of logically independent values x_1, x_2, x_3 take any values, then df = 3. Now $x_1 + x_2 + x_3 = 2$, df = 3 – 1 = 2. 2. If $n \ge 30$, t-distribution approaches to $N(0,1)$, so $N(0,1)$ is used to approximate t-distribution for $n \ge 30$. 3. t-distribution are symmetric about zero, with longer tails than $N(0,1)$. 4. If a random value $X \sim t(df = k)$, then $X^2 \sim F(1, df = k)$.
Exact Binomial Test	**Exact test, no need assumption check.** **Step 1: formulate the hypothesis** # k is the proportion $H_0: p = k \quad v.s. \quad H_1: p \ne k$ (two − tailed) $H_1: p > k$ (right, one − sided, upper − tail) $H_1: p < k$ (left, one − sided, lower − tail) **Step 2: compute the test statistics** $X^* = number\ we\ observed$

	Step 3: compute the p-value or rejection region (RR) $H_1: p \neq k \Rightarrow p-value = 2 * P(X \geq X^*) = 2\big(1 - P(X \leq X^* - 1)\big) = 2(1 - binomcdf(n, k, X^* - 1))$ $H_1: p > k \Rightarrow p-value = P(X \geq X^*) = 1 - P(X \leq X^* - 1) = 1 - binomcdf(n, k, X^* - 1)$ $H_1: p < k \Rightarrow p-value = P(X \leq X^*) = binomcdf(n, k, X^*)$ **Step 4: Make conclusion. If p-value < α, we reject the null hypothesis.**
Two-sample Proportion Z-test	**Assumption:** 1) The number of success and failure must be greater than 5 for both samples. 2) Each sample was randomly taken from the population. **Step 1: formulate the hypothesis** $H_0: p_1 - p_2 = 0 \quad v.s. \quad H_1: p_1 - p_2 \neq 0\ (\text{two} - \text{sided})$ $H_1: p_1 - p_2 > 0\ (\text{right, one} - \text{sided, upper} - \text{tail})$ $H_1: p_1 - p_2 < 0\ (\text{left, one} - \text{sided, lower} - \text{tail})$ **Step 2: Compute the test statistics Z^* based on CLT** $Z^* = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$, where $\hat{p}_1 = \frac{x_1}{n_1}, \hat{p}_2 = \frac{x_2}{n_2}, \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$ **Step 3: Calculate the p-value** $H_1: p_1 - p_2 \neq 0 \Rightarrow p-value = 2 * normalcdf(\lvert Z^* \rvert, 10^{99}, 0, 1)$ $H_1: p_1 - p_2 > 0 \Rightarrow p-value = normalcdf(Z^*, 10^{99}, 0, 1)$ $H_1: p_1 - p_2 < 0 \Rightarrow p-value = normalcdf(Z^*, 10^{99}, 0, 1)$ **Step 4: Make conclusion. If p-value < α, we reject the null hypothesis.**
CI for two-sample Proportion Z-test	(1-α)100% confidence interval for $p_1 - p_2$ is: $(\hat{p}_1 - \hat{p}_2) \pm z_{1-\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ Interpretation: We are (1-α)100% confident that XXXX is between LOWER BOUND and UPPER BOUND.
Two-sample t-test (Paired t-test, General)	**Assumption:** 1) The differences for the matched pairs, follow a normal probability distribution or the sample size (number of pairs) should be larger than or equal to 30. 2) The sample of pairs is a simple random sample from its population. **Step 1: formulate the hypothesis we define $\mu_1 - \mu_2 = d$ or $\mu_2 - \mu_1 = d$** $H_0: d = 0 \quad v.s. \quad H_1: d \neq 0\ (\text{two} - \text{sided})$

	$H_0: d \leq 0 \quad v.s. \quad H_1: d > 0$ (right, one − sided, upper − tail) $H_0: d \geq 0 \quad v.s. \quad H_1: d < 0$ (left, one − sided, lower − tail) **Step 2: Compute the test statistics t^* based on CLT** $t^* = \frac{\bar{d}}{\hat{\sigma}_d/\sqrt{n}} \sim t(df = n-1)$, where $\bar{d} = \frac{\sum d_i}{n}, \hat{\sigma}_d = \sqrt{\frac{\sum_i (d_i - \bar{d})^2}{n-1}}$ **Step 3: Calculate the p-value** $H_1: d \neq 0 \Rightarrow p - value = 2 * tcdf(\lvert t^* \rvert, 10^{99}, n-1)$ $H_1: d > 0 \Rightarrow p - value = tcdf(t^*, 10^{99}, n-1)$ $H_1: d < 0 \Rightarrow p - value = tcdf(t^*, 10^{99}, n-1)$ **Step 4: Make conclusion. If p-value < α, we reject the null hypothesis.**
CI for two-sample t-test (Paired t-test, General)	(1-α)100% confidence interval for the difference between (paired) two means is: $\bar{d} \pm t(1 - \frac{\alpha}{2}, df = n-1)\frac{\hat{\sigma}_d}{\sqrt{n}}$ Interpretation: We are (1-α)100% confident that XXXX is between LOWER BOUND and UPPER BOUND.
Two-sample t-test ($\sigma_1^2 = \sigma_1^2$) ANOVA	**Assumption:** 1) Two samples must be independent 2) Two samples are randomly generated from its population, respectively 3) Two populations have equal variance 4) Each sample size has more than 30 observations **Step 1: formulate the hypothesis** # same with paired t-test $H_0: \mu_1 - \mu_2 = 0 \quad v.s. \quad H_1: \mu_1 - \mu_2 \neq 0$ (two − sided) $H_0: \mu_1 - \mu_2 \leq 0 \quad v.s. \quad H_1: \mu_1 - \mu_2 > 0$ (right, one − sided, upper − tail) $H_0: \mu_1 - \mu_2 \geq 0 \quad v.s. \quad H_1: \mu_1 - \mu_2 < 0$ (left, one − sided, lower − tail) **Step 2: Compute the test statistics t^* based on CLT** $t^* = \frac{\bar{Y}_1 - \bar{Y}_2}{\hat{\sigma}_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(df = n_1 + n_2 - 2)$ where , $\hat{\sigma}_p = \sqrt{\frac{(n_1-1)\hat{\sigma}_1^2 + (n_2-1)\hat{\sigma}_2^2}{n_1+n_2-2}}, \hat{\sigma}_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(Y_{i1} - \bar{Y}_1)^2, \hat{\sigma}_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(Y_{i2} - \bar{Y}_2)^2$ **Step 3: Compute the p-value** $H_1: \mu_1 - \mu_2 \neq 0 \Rightarrow p - value = 2 * tcdf(\lvert t^* \rvert, 10^{99}, n_1 + n_2 - 2)$ $H_1: \mu_1 - \mu_2 > 0 \Rightarrow p - value = tcdf(t^*, 10^{99}, n_1 + n_2 - 2)$ $H_1: \mu_1 - \mu_2 < 0 \Rightarrow p - value = tcdf(t^*, 10^{99}, n_1 + n_2 - 2)$

	Step 4: Make conclusion. If p-value < α, we reject the null hypothesis.
CI for two-sample t-test ($\sigma_1^2 = \sigma_1^2$) ANOVA	(1-α)100% confidence interval for $\mu_1 - \mu_2$ is: $(\bar{Y}_1 - \bar{Y}_2) \pm t\left(1 - \frac{\alpha}{2}, df = n_1 + n_2 - 2\right)\hat{\sigma}_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ Interpretation: We are (1-α)100% confident that XXXX is between LOWER BOUND and UPPER BOUND.
Two-sample t-test ($\sigma_1^2 \neq \sigma_1^2$) NOT ANOVA	**Assumption:** 1) Two samples must be independent 2) Two samples are randomly generated from its population, respectively 3) Each sample size has more than 30 observations **Step 1: formulate the hypothesis** # same with paired t-test $H_0: \mu_1 - \mu_2 = 0 \;\; v.s. \;\; H_1: \mu_1 - \mu_2 \neq 0$ (two − sided) $H_0: \mu_1 - \mu_2 \leq 0 \;\; v.s. \;\; H_1: \mu_1 - \mu_2 > 0$ (right, one − sided, upper − tail) $H_0: \mu_1 - \mu_2 \geq 0 \;\; v.s. \;\; H_1: \mu_1 - \mu_2 < 0$ (left, one − sided, lower − tail) **Step 2: Compute the test statistics t^* based on CLT** $t^* = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} \sim t(df = v)$ where , $\hat{\sigma}_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(Y_{i1} - \bar{Y}_1)^2$, $\hat{\sigma}_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(Y_{i2} - \bar{Y}_2)^2$, $v = \frac{(n_1 - 1)(n_2 - 1)}{(1-c)^2(n_1 - 1) + c^2(n_2 - 1)}$, $c = \frac{\frac{\hat{\sigma}_1^2}{n_1}}{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}$ **Step 3: Compute the p-value** $H_1: \mu_1 - \mu_2 \neq 0 \;\; => \;\; p - value = 2 * tcdf(\lvert t^* \rvert, 10^{99}, v)$ $H_1: \mu_1 - \mu_2 > 0 \;\; => \;\; p - value = tcdf(t^*, 10^{99}, v)$ $H_1: \mu_1 - \mu_2 < 0 \;\; => \;\; p - value = tcdf(t^*, 10^{99}, v)$ **Step 4: Make conclusion. If p-value < α, we reject the null hypothesis.**
CI for two-sample t-test ($\sigma_1^2 \neq \sigma_1^2$) NOT ANOVA	(1-α)100% confidence interval for $\mu_1 - \mu_2$ is: $(\bar{Y}_1 - \bar{Y}_2) \pm t\left(1 - \frac{\alpha}{2}, df = v\right)\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}$ Interpretation: We are (1-α)100% confident that XXXX is between LOWER BOUND and UPPER BOUND.
One-way ANOVA	**Assumption:** 1) Two samples must be independent 2) Two samples are randomly generated from its population, respectively 3) Two populations have equal variance

4) Each sample size has more than 30 observations

The setup of one-way ANOVA

The single-factor mean model is

$Y_{ij} = \mu_1 + \epsilon_{ij}, \quad i = 1,2,\dots,a, \quad j = 1,2,\dots,n_i$

Singe-factor effect model is

$Y_{ij} = \mu + \tau_i + \epsilon_{ij}, \quad i = 1,2,\dots,a, \quad j = 1,2,\dots,n_i$

Facts:

1) $i = 1,2,\dots,a$ denotes the level of the factor/treatment. There are a levels.
2) $j = 1,2,\dots,n_i$ denotes the subject in each factor/treatment level. There are $n_1, n_2, \dots, n_a$ subjects in each level of the factor/treatment.
3) If $n_1 = n_2 = \cdots = n_a$ (every level of the factor/treatment has the same number of subjects), it is called a balanced design. If not, it is an unbalanced design.
4) The mean model and effect model are equivalent, because $\mu_i = \mu + \tau_i$ for $i = 1,2,\dots,a$.
5) ϵ_{ij}'s are model errors. They are unknown and estimated via residuals. The "variance" in Analysis of Variance means "the variance of the residuals".
6) Here we assume ϵ_{ij}'s are independent and identically distributed (IID) and $\epsilon_{ij} \sim N(0, \sigma^2)$ for all i and j.
7) The ANOVA model is called Fixed Effect, because the treatment effects are treated as unknown constants (fixed values) instead of random values. If they are treated as random values, it is called a random effect model.

Notations:

$\hat{\mu} = \bar{Y}_{..}$ (Grand mean), $E(\hat{\mu}) = \mu$
$\hat{\mu}_i = \bar{Y}_{i.}$ (Mean in each level), $E(\hat{\mu}_i) = \mu_i$
$\hat{\tau}_i = \bar{Y}_{i.} - \bar{Y}_{..}$, $E(\hat{\tau}_i) = \tau_i$

Calculation:

Total sample size: $N = n_1 + n_2 + \cdots + n_a$
Overall (Grand) mean (estimate for μ): $\bar{Y}_{..} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{n_i} Y_{ij}$
Treatment means (estimates for $\mu_i's$): $\bar{Y}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}$, $i = 1,2,\dots a$
Treatment effects (estimates for $\tau_i{}'s$): $\bar{Y}_{i.} - \bar{Y}_{..}$, $i = 1,2,\dots a$
$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n_i}\left(Y_{ij} - \bar{Y}_{..}\right)^2$, $df(SS_T) = N - 1$
$SS_{trt} = \sum_{i=1}^{a} n_i(\bar{Y}_{i.} - \bar{Y}_{..})^2$, $df(SStrt) = df(MS_{trt}) = a - 1$, $MS_{trt} = \frac{SS_{trt}}{a-1}$
$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{n_i}\left(Y_{ij} - \bar{Y}_{i.}\right)^2$, $df(SS_E) = df(MS_E) = N - a$, $MS_E = \frac{SS_E}{N-a}$

$\boldsymbol{SS_T = SS_{trt} + SS_E}$

$E(MS_E) = \sigma^2$

$E(MS_{trt}) = \sigma^2 + \frac{\sum_{i=1}^{a} n_i \tau_i^2}{a-1}$

Conducting F-test

Step 1: formulate the hypothesis

$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_a$, $H_1: \mu_i \neq \mu_j$ for at least one pair of (i, j)

or

$H_0: \tau_1 = \tau_2 = \tau_3 = \cdots = \tau_a$, $H_1: \tau_i \neq 0$ for at least one i

Step 2: Compute the test statistics $\boldsymbol{F^*}$

$F^* = \frac{MStrt}{MS_E} \sim F(df_1 = a - 1, df_2 = N - a)$

Here we use the upper tail, so the test statistics here is always greater than 1

df_1 is the degree of freedom of numerator (MS_{trt}) and df_2 is the degree of freedom of denominator (MS_E)

It is okay to switch the numerator and the denominator, but the degrees of freedom and p-value of the F distribution should also be switched.

Step 3: Compute the p-value

$p - value = Fcdf(F^*, 10^{99}, df_1 = a - 1, df_2 = N - a)$

Step 4: Make conclusion. If p-value < α, we reject the null hypothesis.

ANOVA table

Source of Variation	Sum of Squares	df	Mean Square	$F-$score
Between Treatments	SS_{trt}	$a-1$	$MS_{trt} = \frac{SS_{trt}}{a-1}$	$F^* = \frac{MS_{trt}}{MS_E}$
Error(Within Treatments)	SS_E	$N-a$	$MS_E = \frac{SS_E}{N-a}$	
Total	SS_T	$N-1$		

Model Assumption Check

1. The sample (the design) is a random sample.

2. The errors are normally distributed

Method 1: QQ plot (quantile-quantile plot).

Method 2: Shapiro-Wilk normality test.

H_0: The data are normally distributed, H_1: The data are NOT normally distributed
If the p-value < α, we reject the null hypothesis.

3. The errors are independently distributed
Method 1: Plot of residuals in time series (if the order of the runs is known)
Method 2: Plot of residuals $\hat{\epsilon}_{ij}$'s versus fitted values $\bar{Y}_i$'s.

4. The errors have equal variances
Method 1: Grouped boxplot.
If the widths of the boxes are similar, then we conclude they have equal variance. (NOT accurate)
Method 2: Tests for equal variance
H_0: All groups have equal variances ($\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2$), H_1: All groups have UNEQUAL variances
Levene's test: Assess the equality of variances for a variable calculated for two or more samples (groups/levels).
1) $d_{ij} = |Y_{ij} - \bar{Y}_{i.}|$, for $i = 1,2,\dots,a;\ \ j = 1,2,\dots,n_i$
2) Run on-way ANOVA test on d_{ij}
Brown-Forsythe test: Use median instead of mean in Levene's test.
$d_{ij} = \left|Y_{ij} - \tilde{Y}_{i.}\right|$, for $i = 1,2,\dots,a;\ \ j = 1,2,\dots,n_i$
Bartlett's test: Derived from the likelihood ratio test under normality assumption. (NOT required for this class)
Brown-Forsythe test should be preferred, it is the "median" version of Levene's test. If the data are skewed, Brown-Forsythe is more robust; if the data are symmetric (not skewed), Brown-Forsythe test and Levene's test have similar performance.

Comparisons among treatment level means
$H_0: \mu_i = \mu_j$, $H_1: \mu_i \neq \mu_j$ or $H_0: \mu_{i_1} + \mu_{i_2} = \mu_{j_1} + \mu_{j_2}$, $H_1: \mu_{i_1} + \mu_{i_2} \neq \mu_{j_1} + \mu_{j_2}$

We can use contrast to perform this above
$\Gamma = \sum_{i=1}^{a} c_i \mu_i$
$H_0: \sum_{i=1}^{a} c_i \mu_i = 0$, $H_1: \sum_{i=1}^{a} c_i \mu_i \neq 0$
Point estimate: $C = \sum_{i=1}^{a} c_i \bar{Y}_{i.}$
Sample variance: $Var(C) = MS_E \sum_{i=1}^{a} \frac{c_i^2}{n_i}$
Standard error: $s(C) = \sqrt{Var(C)} = \sqrt{MS_E \sum_{i=1}^{a} \frac{c_i^2}{n_i}}$
Test statistic: $F^* = \frac{C^2}{Var(C)} = \frac{\left(\sum_{i=1}^{a} c_i \bar{Y}_{i.}\right)^2}{MS_E \sum_{i=1}^{a} \frac{c_i^2}{n_i}} \sim F(df_1 = 1,\ df_2 = N - a)$
We reject the null if p-value < α.

	(1-α)100% confidence interval on the contrast $\Gamma = \sum_{i=1}^{a} c_i\mu_i$ is: $C \pm t\left(1-\frac{\alpha}{2}, df = N-a\right)s(C)$ **Scheffé Method for Comparing all contrasts (simultaneous method)** $s(C) = \sqrt{Var(C)} = \sqrt{MS_E \sum_{i=1}^{a} \frac{c_i^2}{n_i}}$ We reject the null hypothesis $H_0: \sum_{i=1}^{a} c_i\mu_i = 0$ if $\lvert C\rvert > s(C)\sqrt{(a-1)F(1-\alpha, df_1 = a-1, df_2 = N-a)}$ Simultaneous confidence intervals: $C \pm s(C)\sqrt{(a-1)F(1-\alpha, df_1 = a-1, df_2 = N-a)}$ **Comparing pairs of treatment means** $H_0: \mu_i - \mu_j = 0$, $H_1: \mu_i - \mu_j \neq 0$ for all $i \neq j$ **Tukey HSD:** We reject null hypothesis if: $\lvert\bar{Y}_{i.} - \bar{Y}_{j.}\rvert > \frac{q_\alpha(a, df=N-a)}{\sqrt{2}}\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$ Tukey's (1-α)100% confidence interval for $\mu_i - \mu_j$ is $(\bar{Y}_{i.} - \bar{Y}_{j.}) \pm \frac{q_\alpha(a, df=N-a)}{\sqrt{2}}\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$ **Fisher Least Significant Difference (LSD):** We reject null hypothesis if: $\lvert\bar{Y}_{i.} - \bar{Y}_{j.}\rvert > t\left(1-\frac{\alpha}{2}, df = N-a\right)\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$ Fisher's (1-α)100% confidence interval for $\mu_i - \mu_j$ is $(\bar{Y}_{i.} - \bar{Y}_{j.}) \pm t\left(1-\frac{\alpha}{2}, df = N-a\right)\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$ # same as CI on the difference in any two treatment means $\mu_i - \mu_j$
CI for one-way ANOVA (pairwise)	(1-α)100% confidence interval for the i-th treatment mean μ_i is: $\bar{Y}_{i.} \pm t\left(1-\frac{\alpha}{2}, df = N-a\right)\sqrt{\frac{MS_E}{n_i}}$ (1-α)100% confidence interval on the difference in any two treatment means $\mu_i - \mu_j$ is: $\left(\bar{Y}_{i.} - \bar{Y}_{j.}\right) \pm t\left(1-\frac{\alpha}{2}, df = N-a\right)\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$