Chapter 1

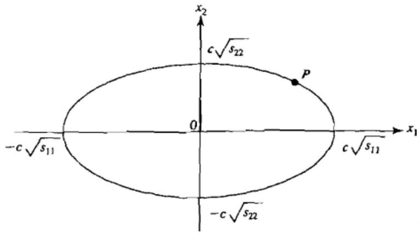


Figure 1.21 The ellipse of constant statistical distance $d^2(O,P) = x_1^2/s_{11} + x_2^2/s_{22} = c^2$.

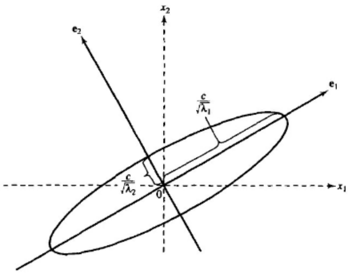


Figure 2.6 Points a constant distance c from the origin ($p = 2, 1 \le \lambda_1 < \lambda_2$).

If $p > 2$, the points $\mathbf{x}' = [x_1, x_2, \ldots, x_p]$ a constant distance $c = \sqrt{\mathbf{x}'\mathbf{A}\mathbf{x}}$ from the origin lie on hyperellipsoids $c^2 = \lambda_1(\mathbf{x}'\mathbf{e}_1)^2 + \cdots + \lambda_p(\mathbf{x}'\mathbf{e}_p)^2$, whose axes are given by the eigenvectors of $\mathbf{A}$. The half-length in the direction $\mathbf{e}_i$ is equal to $c/\sqrt{\lambda_i}$, $i = 1, 2, \ldots, p$, where $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of $\mathbf{A}$.

Chapter 2A

Vectors and Matrices

Angle θ between X and Y :

$$\cos(\theta) = \frac{x'y}{L_x L_y} \Rightarrow \theta = \cos^{-1}\left(\frac{x'y}{L_x L_y}\right)$$

Projection of X on Y = proj(X, Y) :

$$\text{proj}(X,Y) = \frac{X'Y}{Y'Y} Y ,$$

Length of the projection :

$$\frac{|X'Y|}{L_Y} = L_X \left|\frac{X'Y}{L_X L_Y}\right| = L_X |\cos\theta|$$

Gram-Schmidt Process :

$$u_1 = x_1$$

$$u_2 = x_2 - \frac{x_2' u_1}{u_1' u_1} u_1$$

$$\vdots$$

$$u_k = x_k - \frac{x_k' u_1}{u_1' u_1} - \cdots - \frac{x_k' u_{k-1}}{u_{k-1}' u_{k-1}} u_{k-1}$$

Quadratic Form :

$$Q(X) = x'Ax = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}^2 x_i x_j$$

Any symmetric square matrix can be reconstructed from its eigenvalues and eigenvectors. The particular expression reveals the relative importance of each pair according to the relative size of the eigenvalue and the direction of the eigenvector.

Spectral Decomposition

$$A = \sum_{i=1}^{k} \lambda_i e_i e_i' ,$$ λ_i = eigenvalues , e_i = corresponding eigenvector

The ideas that lead to the spectral decomposition can be extended to provide a decomposition for a rectangular, rather than a square, matrix. If $\mathbf{A}$ is a rectangular matrix, then the vectors in the expansion of $\mathbf{A}$ are the eigenvectors of the square matrices $\mathbf{AA}'$ and $\mathbf{A}'\mathbf{A}$.

Singular-Value Decomposition :

$$A_{m\times k} = U\Lambda V'$$

$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots \lambda_r)$

U = $m\times m$ orthogonal matrix

V = $k\times k$ orthogonal matrix

$$A = \sum_{i=1}^{r} \lambda_i u_i u_i' = U_r \Lambda_r V_r$$

Chapter 2

Matrix Algebra and Random Vectors

Spectral Decomposition : an expansion for symmetric matrices

$A = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \cdots \lambda_k e_k e_k'$, e_i = normalized eigenvectors

* e_i's are the normalized solutions of the equations $Ae_i = \lambda_i e_i$

* Symmetric matrices are positive definite matrices iif $\lambda_i > 0$

nonnegative // // $\lambda_i \geq 0$

$$(\text{distance})^2 = x'Ax = \sum \lambda_i x' e_i e_i' x$$

Square-Root Matrix :

Let $A = \sum_{i=1}^{k} \lambda_i e_i e_i' = P\Lambda P'$, where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_k)$

$P = [e_1 \ e_2 \ \cdots \ e_k]$

Then $A^{-1} = \sum_{i=1}^{k} \frac{1}{\lambda_i} e_i e_i' = P\Lambda^{-1}P'$

$$A^{\frac{1}{2}} = \sum_{i=1}^{k} \sqrt{\lambda_i}\, e_i e_i' = P\Lambda^{\frac{1}{2}}P'$$

square root of A

$$\boldsymbol{\Sigma} = E(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})' = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix}$$

$$\mathbf{V}^{1/2} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{pp}} \end{bmatrix}$$

$$\boldsymbol{\rho} = \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} \\ \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{22}}} & \cdots & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} & \cdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{pp}}} \end{bmatrix} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \rho_{2p} & \cdots & 1 \end{bmatrix}$$

$$\mathbf{V}^{1/2}\boldsymbol{\rho}\mathbf{V}^{1/2} = \boldsymbol{\Sigma}$$

$$\boldsymbol{\rho} = (\mathbf{V}^{1/2})^{-1}\boldsymbol{\Sigma}(\mathbf{V}^{1/2})^{-1}$$

The Mean Vector and Covariance Matrix for Linear Combinations of Random Variables

$$\begin{aligned} Cov(aX_1, bX_2) &= E(aX_1 - a\mu_1)(bX_2 - b\mu_2) \\ &= ab\,E(X_1-\mu_1)(X_2-\mu_2) \\ &= ab\,Cov(X_1, X_2) = ab\,\sigma_{12} \end{aligned}$$

$$\begin{aligned} Var(aX_1 + bX_2) &= E[(aX_1+bX_2) - (a\mu_1+b\mu_2)]^2 \\ &= E[a(X_1-\mu_1) + b(X_2-\mu_2)]^2 \\ &= E[a^2(X_1-\mu_1)^2 + b^2(X_2-\mu_2)^2 + 2ab(X_1-\mu_1)(X_2-\mu_2)] \\ &= a^2 Var(X_1) + b^2 Var(X_2) + 2ab\,Cov(X_1, X_2) \\ &= a^2\sigma_{11} + b^2\sigma_{22} + 2ab\,\sigma_{12} \\ &= Var(c'X) \\ &= c'\Sigma c \end{aligned}$$

$$\mu_Z = E(Z) = E(CX) = C\mu_X$$

$$\Sigma_Z = Cov(Z) = cov(CX) = C\Sigma_X C'$$

Chapter 3 Sample Geometry and Random Sampling

Squared Lengths of the deviation vectors

$$\underbrace{L_{d_i}^2}_{(\text{Length})^2} = d_i' d_i = \underbrace{\sum_{j=1}^{n} (x_{ji} - \bar{x}_i)^2}_{\text{sum of squared deviations}}$$

* note that the squared length L^2 is proportional to the variance of the measurements on the i^{th} variable. Equivalently, the length is proportional to the standard deviation. Longer vectors represent more variability than shorter vectors.

For any two deviation vectors d_i and d_k,

$d_i' d_k = \sum_{j=1}^{n} (x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)$, and let θ_{ik} be the angle between d_i and d_k,

$d_i' d_k = L_{d_i} L_{d_k} \cos(\theta_{ik}) = \sqrt{\sum_{j=1}^{n}(x_{ji}-\bar{x}_i)^2} \sqrt{\sum_{j=1}^{n}(x_{jk}-\bar{x}_k)^2} \cdot \cos(\theta_{ik})$, then we notice that

$$\frac{d_i' d_k}{L_{d_i} L_{d_k}} = \frac{\sum_{j=1}^{n}(x_{ji}-\bar{x}_i)(x_{jk}-\bar{x}_k)}{\sqrt{\sum_{j=1}^{n}(x_{ji}-\bar{x}_i)^2 \cdot \sum_{j=1}^{n}(x_{jk}-\bar{x}_k)^2}} = \cos(\theta_{ik}) = r_{ik} = Cor(i, k)$$

$\cos(0°) = 1$, $\cos(90°) = 0$, ... 등등 Cosine의 각도를 생각해보면 당연한 말...! 씬기해...!

Generalized Variance :

- Let S be a variance-covariance matrix, then $|S|$ = Generalized Variance

Chapter 4 Multivariate Normal Distribution

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} e^{-(X-\mu)'\Sigma^{-1}(X-\mu)/2}, \quad -\infty < x < \infty, \quad i = 1, 2, \dots, p \sim N_p(\mu, \Sigma)$$

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \exp\left\{-\frac{1}{2(1-\rho_{12}^2)}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\}$$

- We need to know that the constant $c^2 = \chi^2_p(\alpha)$, where $\chi^2_p(\alpha)$ is the upper $(100\alpha)^{th}$ percentile of a chi-square distribution with p degrees of freedom, which means contours contain $(1-\alpha)\times 100\%$ of the probability

$$(x-\mu)'\Sigma^{-1}(x-\mu) \le \chi^2_p(\alpha)$$

Theorem 2

Suppose that

$$\mathbf{X} = \left[\frac{\mathbf{X}_{(1)}}{\mathbf{X}_{(2)}}\right] \sim N_{q_1+q_2}\left(\left[\frac{\boldsymbol{\mu}_{(1)}}{\boldsymbol{\mu}_{(2)}}\right], \left[\begin{array}{c|c} \Sigma_{(11)} & \Sigma_{(12)} \\ \hline \Sigma_{(21)} & \Sigma_{(22)} \end{array}\right]\right).$$

Then, the following properties hold:

marginally normal

1. $\mathbf{X}_{(1)}$ *and* $\mathbf{X}_{(2)}$ *are independent if and only if* $\Sigma_{(12)} = \mathbf{0}_{q_1 \times q_2}$.
2. $\mathbf{X}_{(j)} \sim N_{q_j}(\boldsymbol{\mu}_{(j)}, \Sigma_{(jj)})$ *for* $j = 1, 2$.
3. $\mathbf{X}_{(1)}|\mathbf{X}_{(2)} = \mathbf{x}_{(2)} \sim N_{q_1}(\boldsymbol{\mu}_{1|2}, \Sigma_{1|2})$
   where $\boldsymbol{\mu}_{1|2} = \boldsymbol{\mu}_{(1)} + \Sigma_{(12)}\Sigma_{(22)}^{-1}(\mathbf{x}_{(2)} - \boldsymbol{\mu}_{(2)})$ *and*
   $\Sigma_{1|2} = \Sigma_{(11)} - \Sigma_{(12)}\Sigma_{(22)}^{-1}\Sigma_{(21)}$.

Maximum Likelihood Estimation of μ and Σ

- Given a pxp symmetric positive definite matrix B and a scalar $b > 0$, it follows that,

$$\frac{1}{|\Sigma|^b} e^{-tr(\Sigma^{-1}B)/2} \le \frac{1}{|B|^b}(2b)^{pb} e^{-bp},$$

for all positive definite Σ, with equality holding only for $\Sigma = (1/2b)B$.

Constant Probability Density Contour $= \{$all x such that $(x-\mu)'\Sigma^{-1}(x-\mu) = c^2\}$
$=$ surface of an ellipsoid centered at μ.

$\pm c\sqrt{\lambda_i}e_i$: points of the tip of the axes

For the multivariate normal situation, it is worth emphasizing the following:

1. All conditional distributions are (multivariate) normal.
2. The conditional mean is of the form

$$\mu_1 + \beta_{1,q+1}(X_{q+1} - \mu_{q+1}) + \cdots + \beta_{1,p}(X_p - \mu_p)$$
$$\vdots$$
$$\mu_q + \beta_{q,q+1}(X_{q+1} - \mu_{q+1}) + \cdots + \beta_{q,p}(X_p - \mu_p)$$

where the β's are defined by $\Sigma_{12}\Sigma_{22}^{-1} = \begin{bmatrix} \beta_{1,q+1} & \beta_{1,q+2} & \cdots & \beta_{1,p} \\ \beta_{2,q+1} & \beta_{2,q+2} & \cdots & \beta_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q,q+1} & \beta_{q,q+2} & \cdots & \beta_{q,p} \end{bmatrix}$

3. The conditional covariance, $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, does not depend upon the value(s) of the conditioning variable(s).

Maximum Likelihood Estimates :

- selecting the parameter values that maximize the joint density evaluated at the observations.

- Consider the joint density function

$$\left\{ \begin{matrix} \text{Joint density} \\ \text{of } \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n \end{matrix} \right\} = \prod_{j=1}^{n} \left\{ \frac{1}{(2\pi)^{p/2}|\mathbf{\Sigma}|^{1/2}} e^{-(\mathbf{x}_j-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}_j-\boldsymbol{\mu})/2} \right\}$$
$$= \frac{1}{(2\pi)^{np/2}} \frac{1}{|\mathbf{\Sigma}|^{n/2}} e^{-\sum_{j=1}^{n}(\mathbf{x}_j-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{x}_j-\boldsymbol{\mu})/2}$$

, and now,

the exponent in there can be simplified.

$$\sum_{j=1}^{n}(x_j-\mu)'\Sigma^{-1}(x_j-\mu) = tr\left[\Sigma^{-1}\left(\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})' + n(\bar{x}-\mu)(\bar{x}-\mu)'\right)\right]$$
$$= tr\left[\Sigma^{-1}\left(\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})'\right)\right] + n(\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu)$$

Maximum Likelihood Estimation of μ and Σ

- Given a $p\times p$ symmetric positive definite matrix B and a scalar $b>0$, it follows that,

$$\frac{1}{|\Sigma|^b} e^{-tr(\Sigma^{-1}B)/2} \le \frac{1}{|B|^b}(2b)^{pb} e^{-bp},$$

for all positive definite Σ, with equality holding only for $\Sigma = (1/2b)B$.

Chapter 5

Inferences About a Mean Vector

$$\bar{x} - t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}} \le \mu_0 \le \bar{x} + t_{n-1}(\alpha/2)\frac{s}{\sqrt{n}}$$

* $t^2 \sim F$

=> This is a random interval because the endpoints depend upon the random variables $\bar{X}$ and S.

- When the test is multivariate, which X is a $n \times p$ matrix and $\bar{X}$ is a $p \times 1$ matrix,

$T^2 = (\bar{X} - \mu_0)' \left(\frac{1}{n}S\right)^{-1} (\bar{X} - \mu_0) = n(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0)$, where

$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$, $S = \frac{1}{n-1}\sum_{j=1}^{n} (X_j - \bar{X})(X_j - \bar{X})'$, and this T^2 is called Hotelling's T^2.

- T^2 is distributed as $\frac{(n-1)p}{(n-p)} F_{p,n-p}$, where $F_{p,n-p}$ denotes a random variable with an F-distribution p and $n-p$ degrees of freedom.

=> $\alpha = P\left[T^2 > \frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)\right] = P\left[n(\bar{X} - \mu)' S^{-1} (\bar{X} - \mu) > \frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)\right]$

1. Compute $\bar{X}$, S^{-1}
2. Compute T^2
3. Calculate $\frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)$, and Compare it with T^2
4. H_0 will be rejected if one or more of the component means, or some combination of means, differs too much from the hypothesized values.

* One feature of the T^2-statistic is that it is invariant under changes in the units of measurements for X of the form,

$Y = CX + d$, C : nonsingular. Premultiplication of the centered and scaled quantities by any nonsingular matrix will be invariant.

Hotelling's T^2 and Likelihood Ratio Tests

- Likelihood ratio tests have several optimal properties for reasonably large samples, and they are particularly convenient for hypotheses formulated in terms of multivariate normal parameters.

$\hat{\Sigma}_0 = \frac{1}{n}\sum_{j=1}^{n} (X_j - \mu_0)(X_j - \mu_0)'$

- Likelihood Ratio $= \Lambda = \dfrac{\max_{\Sigma} L(\mu_0, \Sigma)}{\max_{\mu,\Sigma} L(\mu, \Sigma)} = \left(\dfrac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}\right)^{\frac{n}{2}}$ or => $\Lambda^{\frac{2}{n}} = \dfrac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}$

, and this is called Wilks' lambda.

- H_0 is rejected if the likelihood ratio is too small.

* note that the likelihood ratio test statistic is a power of the ratio of generalized variances.

Theorem : * Result 5.1 pdf 239

- Let $X_1, X_2, \ldots, X_n$ be a random sample from an $N_p(\mu, \Sigma)$ population. Then the test based on T^2 is equivalent to the likelihood ratio test of $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ because,

$$\Lambda^{\frac{2}{n}} = \left(1 + \frac{T^2}{(n-1)}\right)^{-1} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}$$

Confidence Regions and Simultaneous Comparisons of Component Means

- $100(1-\alpha)\%$ confidence region for the mean of a p-dimensional normal distribution is the ellipsoid determined by all μ such that,

$$n(\bar{x}-\mu)'S^{-1}(\bar{x}-\mu) \leq \frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha) = c^2$$

- To determine whether any given μ_0 lies within the confidence region, we need to compute the generalized squared distance $n(\bar{x}-\mu_0)'S^{-1}(\bar{x}-\mu_0)$ and compare it with $[p(n-1)/(n-p)]F_{p,n-p}(\alpha)$

$$\Rightarrow \quad \frac{p(n-1)}{(n-p)} F_{p,n-p}(\alpha) = c^2$$

$$\Rightarrow \quad \sqrt{\lambda_i}\, c/\sqrt{n} = \sqrt{\lambda_i}\sqrt{p(n-1)F_{p,n-p}(\alpha)/n(n-p)}$$

half-lengths of the major & minor axes

$$\bar{x} \pm \sqrt{\lambda_i}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p}(\alpha)} \cdot e_i$$

the axes of confidence ellipsoids

* the ratio of the λ_i's will help identify relative amounts of elongation along pairs of axes

- Simultaneous Confidence Intervals

It is convenient to refer to the simultaneous intervals of Result 5.3 as T^2-*intervals*, since the coverage probability is determined by the distribution of T^2. The successive choices $\mathbf{a}' = [1, 0, \ldots, 0]$, $\mathbf{a}' = [0, 1, \ldots, 0]$, and so on through $\mathbf{a}' = [0, 0, \ldots, 1]$ for the T^2-intervals allow us to conclude that

$$\begin{aligned}
\bar{x}_1 - \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{11}}{n}} &\leq \mu_1 \leq \bar{x}_1 + \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{11}}{n}} \\
\bar{x}_2 - \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{22}}{n}} &\leq \mu_2 \leq \bar{x}_2 + \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{22}}{n}} \\
&\vdots \\
\bar{x}_p - \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{pp}}{n}} &\leq \mu_p \leq \bar{x}_p + \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{pp}}{n}}
\end{aligned} \qquad (5\text{-}24)$$

- Bonferroni Method

Exercise 5.6),

$$P[\text{all } C_i \text{ true}] = 1 - P[\text{at least one } C_i \text{ false}]$$
$$\geq 1 - \sum_{i=1}^{m} P(C_i \text{ false}) = 1 - \sum_{i=1}^{m} (1 - P(C_i \text{ true}))$$
$$= 1 - (\alpha_1 + \alpha_2 + \cdots + \alpha_m) \qquad (5\text{-}28)$$

Inequality (5-28), a special case of the Bonferroni inequality, allows an investigator to control the overall error rate $\alpha_1 + \alpha_2 + \cdots + \alpha_m$, regardless of the correlation structure behind the confidence statements. There is also the flexibility of controlling the error rate for a group of important statements and balancing it by another choice for the less important statements.

Let us develop simultaneous interval estimates for the restricted set consisting of the components μ_i of $\boldsymbol{\mu}$. Lacking information on the relative importance of these components, we consider the individual t-intervals

applied

$$\bar{x}_i \pm t_{n-1}\left(\frac{\alpha_i}{2}\right)\sqrt{\frac{s_{ii}}{n}} \qquad i = 1, 2, \ldots, m$$

with $\alpha_i = \alpha/m$. Since $P[\bar{X}_i \pm t_{n-1}(\alpha/2m)\sqrt{s_{ii}/n}$ contains $\mu_i] = 1 - \alpha/m$, $i = 1, 2, \ldots, m$, we have, from (5-28),

$$P\left[\bar{X}_i \pm t_{n-1}\left(\frac{\alpha}{2m}\right)\sqrt{\frac{s_{ii}}{n}} \text{ contains } \mu_i, \text{ all } i\right] \geq 1 - \underbrace{\left(\frac{\alpha}{m} + \frac{\alpha}{m} + \cdots + \frac{\alpha}{m}\right)}_{m \text{ terms}}$$
$$= 1 - \alpha$$

Therefore, with an overall confidence level greater than or equal to $1 - \alpha$, we can make the following $m = p$ statements:

$$\bar{x}_1 - t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{11}}{n}} \leq \mu_1 \leq \bar{x}_1 + t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{11}}{n}}$$
$$\bar{x}_2 - t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{22}}{n}} \leq \mu_2 \leq \bar{x}_2 + t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{22}}{n}} \qquad (5\text{-}29)$$
$$\vdots \qquad\qquad \vdots$$
$$\bar{x}_p - t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{pp}}{n}} \leq \mu_p \leq \bar{x}_p + t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{s_{pp}}{n}}$$

✻ If we are interested only in the component means, the Bonferroni intervals provide more precise estimates than the T^2-intervals. On the other hand, the 95% confidence region for μ gives the plausible values for the pairs (μ_1, μ_2) when the correlation between the measured variables is taken into account.

$$\frac{\text{Length of Bonferroni Interval}}{\text{Length of } T^2\text{-interval}} = \frac{t_{n-1}\left(\frac{\alpha}{2m}\right)}{\sqrt{\frac{p(n-1)}{n-p} F_{p,n-p}(\alpha)}}$$

Large Sample Inference about a Population Mean Vector

✻ All large-sample inferences about μ are based on a χ^2-distribution

$$P[n(\bar{x}-\mu)'S^{-1}(\bar{x}-\mu) \leq \chi^2_p(\alpha)] = 1-\alpha$$

✻ $\frac{(n-1)p}{(n-p)} F_{p,n-p}(\alpha)$ and $\chi^2_p(\alpha)$ are approximately equal for n large relative to p.

Result 5.5. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be a random sample from a population with mean $\boldsymbol{\mu}$ and positive definite covariance $\boldsymbol{\Sigma}$. If $n - p$ is large,

$$\mathbf{a}'\bar{\mathbf{X}} \pm \sqrt{\chi^2_p(\alpha)}\sqrt{\frac{\mathbf{a}'\mathbf{S}\mathbf{a}}{n}}$$

will contain $\mathbf{a}'\boldsymbol{\mu}$, for every $\mathbf{a}$, with probability approximately $1 - \alpha$. Consequently, we can make the $100(1 - \alpha)\%$ simultaneous confidence statements

$$\bar{x}_1 \pm \sqrt{\chi^2_p(\alpha)}\sqrt{\frac{s_{11}}{n}} \quad \text{contains } \mu_1$$
$$\bar{x}_2 \pm \sqrt{\chi^2_p(\alpha)}\sqrt{\frac{s_{22}}{n}} \quad \text{contains } \mu_2$$
$$\vdots \qquad\qquad \vdots$$
$$\bar{x}_p \pm \sqrt{\chi^2_p(\alpha)}\sqrt{\frac{s_{pp}}{n}} \quad \text{contains } \mu_p$$

and, in addition, for all pairs (μ_i, μ_k), $i, k = 1, 2, \ldots, p$, the sample mean-centered ellipses

$$n[\bar{x}_i - \mu_i, \ \ \bar{x}_k - \mu_k]\begin{bmatrix} s_{ii} & s_{ik} \\ s_{ik} & s_{kk} \end{bmatrix}^{-1}\begin{bmatrix} \bar{x}_i - \mu_i \\ \bar{x}_k - \mu_k \end{bmatrix} \leq \chi^2_p(\alpha) \quad \text{contain} \quad (\mu_i, \mu_k)$$

Chapter 6

Comparisons of Several Multivariate Means

Paired Comparisons and a Repeated Measures Design

Paired Comparisons :

- Let $\bar{D}$ denotes the difference vector of means and δ is the population parameter of $\bar{D}$.

  Inferences about the vector of mean differences δ can be based on a T^2-statistic

  $T^2 = n(\bar{D}-\delta)' S_d^{-1}(\bar{D}-\delta)$, where $\bar{\mathbf{D}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{D}_j$ and $\mathbf{S}_d = \frac{1}{n-1}\sum_{j=1}^{n}(\mathbf{D}_j-\bar{\mathbf{D}})(\mathbf{D}_j-\bar{\mathbf{D}})'$, and T^2 is distributed as an $\frac{(n-1)p}{(n-p)}F_{p,n-p}$.

  If n and n-p are both large, T^2 is approximately distributed as a χ^2 random variable, regardless of the form of the underlying population of differences.

- Having the same covariance matrix is very important, and with Σ_1 and Σ_2, we are able to compute the common covariance Σ, using the below equation.

$$\mathbf{S}_{\text{pooled}} = \frac{\sum_{j=1}^{n_1}(\mathbf{x}_{1j}-\bar{\mathbf{x}}_1)(\mathbf{x}_{1j}-\bar{\mathbf{x}}_1)' + \sum_{j=1}^{n_2}(\mathbf{x}_{2j}-\bar{\mathbf{x}}_2)(\mathbf{x}_{2j}-\bar{\mathbf{x}}_2)'}{n_1+n_2-2} = \frac{n_1-1}{n_1+n_2-2}\mathbf{S}_1 + \frac{n_2-1}{n_1+n_2-2}\mathbf{S}_2 \tag{6-21}$$

Since the independence assumption in (6-19) implies that $\bar{\mathbf{X}}_1$ and $\bar{\mathbf{X}}_2$ are independent and thus $\text{Cov}(\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2) = \mathbf{0}$ (see Result 4.5), by (3-9), it follows that

$$\text{Cov}(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) = \text{Cov}(\bar{\mathbf{X}}_1) + \text{Cov}(\bar{\mathbf{X}}_2) = \frac{1}{n_1}\boldsymbol{\Sigma} + \frac{1}{n_2}\boldsymbol{\Sigma} = \left(\frac{1}{n_1}+\frac{1}{n_2}\right)\boldsymbol{\Sigma} \tag{6-22}$$

Because $\mathbf{S}_{\text{pooled}}$ estimates $\boldsymbol{\Sigma}$, we see that

$$\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_{\text{pooled}}$$

is an estimator of $\text{Cov}(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2)$.

The likelihood ratio test of

$$H_0: \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \boldsymbol{\delta}_0$$

is based on the square of the statistical distance, T^2, and is given by (see [1]). Reject H_0 if

$$T^2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \boldsymbol{\delta}_0)'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_{\text{pooled}}\right]^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \boldsymbol{\delta}_0) > c^2 \tag{6-23}$$

Result 6.2. If $\mathbf{X}_{11}, \mathbf{X}_{12}, \ldots, \mathbf{X}_{1n_1}$ is a random sample of size n_1 from $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ and $\mathbf{X}_{21}, \mathbf{X}_{22}, \ldots, \mathbf{X}_{2n_2}$ is an independent random sample of size n_2 from $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$, then

$$T^2 = [\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)]'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_{\text{pooled}}\right]^{-1}[\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)]$$

is distributed as

$$\frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)}F_{p,n_1+n_2-p-1}$$

Consequently,

$$P\left[(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2))'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_{\text{pooled}}\right]^{-1}(\bar{\mathbf{X}}_1 - \quad \sqrt{\lambda_i}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)c^2} = \sqrt{\lambda_i}\,\sqrt{.25}\right] \tag{6-24}$$

where

$$c^2 = \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)}F_{p,n_1+n_2-p-1}(\alpha)$$

the lengths of the major/minor axes of an ellipse

$\sqrt{\lambda_i}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)c^2}$

$$T^2 = (\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2))'\left[\frac{1}{n_1}S_1 + \frac{1}{n_2}S_2\right]^{-1}(\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2))$$

$$T^2 = \frac{\nu p}{\nu-p+1}F_{p,\nu-p+1}$$

$$\nu = \frac{p+p^2}{\sum_{i=1}^{2}\frac{1}{n_i}\left\{\text{tr}\left[\left(\frac{1}{n_i}\mathbf{S}_i\left(\frac{1}{n_1}\mathbf{S}_1+\frac{1}{n_2}\mathbf{S}_2\right)^{-1}\right)^2\right] + \left(\text{tr}\left[\frac{1}{n_i}\mathbf{S}_i\left(\frac{1}{n_1}\mathbf{S}_1+\frac{1}{n_2}\mathbf{S}_2\right)^{-1}\right]\right)^2\right\}}$$

Simultaneous Confidence Intervals

Result 6.3. Let $c^2 = [(n_1+n_2-2)p/(n_1+n_2-p-1)]F_{p,n_1+n_2-p-1}(\alpha)$. With probability $1-\alpha$.

$$\mathbf{a}'(\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2) \pm c\sqrt{\mathbf{a}'\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathbf{S}_{\text{pooled}}\mathbf{a}}$$

will cover $\mathbf{a}'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$ for all $\mathbf{a}$. In particular $\mu_{1i} - \mu_{2i}$ will be covered by

$$(\bar{X}_{1i} - \bar{X}_{2i}) \pm c\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)s_{ii,\text{pooled}}} \quad \text{for } i = 1, 2, \ldots, p$$

The Bonferroni $100(1-\alpha)\%$ *simultaneous confidence intervals* for the p population mean differences are

$$\mu_{1i} - \mu_{2i}: \quad (\bar{x}_{1i} - \bar{x}_{2i}) \pm t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right)\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)s_{ii,\text{pooled}}}$$

where $t_{n_1+n_2-2}(\alpha/2p)$ is the upper $100(\alpha/2p)$th percentile of a t-distribution with n_1+n_2-2 d.f.

One-Way Univariate ANOVA

ANOVA Table for Comparing Univariate Population Means

Source of variation	Sum of squares (SS)	Degrees of freedom (d.f.)
Treatments	$SS_{tr} = \sum_{\ell=1}^{g} n_\ell(\bar{x}_\ell - \bar{x})^2$	$g - 1$
Residual (error)	$SS_{res} = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(x_{\ell j} - \bar{x}_\ell)^2$	$\sum_{\ell=1}^{g} n_\ell - g$
Total (corrected for the mean)	$SS_{cor} = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(x_{\ell j} - \bar{x})^2$	$\sum_{\ell=1}^{g} n_\ell - 1$

One-Way MANOVA

MANOVA Table for Comparing Population Mean Vectors

Source of variation	Matrix of sum of squares and cross products (SSP)	Degrees of freedom (d.f.)
Treatment	$\mathbf{B} = \sum_{\ell=1}^{g} n_\ell(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$	$g - 1$
Residual (Error)	$\mathbf{W} = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$	$\sum_{\ell=1}^{g} n_\ell - g$
Total (corrected for the mean)	$\mathbf{B} + \mathbf{W} = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'$	$\sum_{\ell=1}^{g} n_\ell - 1$

Testing parameter

Wishart densities. (See [1].)

One test of $H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \mathbf{0}$ involves generalized variances. We reject H_0 if the ratio of generalized variances

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|} = \frac{\left|\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'\right|}{\left|\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'\right|} \tag{6-42}$$

is too small. The quantity $\Lambda^* = |\mathbf{W}|/|\mathbf{B} + \mathbf{W}|$, proposed originally by Wilks

Bartlett (see [4]) has shown that if H_0 is true and $\Sigma n_\ell = n$ is large,

$$-\left(n - 1 - \frac{(p+g)}{2}\right)\ln \Lambda^* = -\left(n - 1 - \frac{(p+g)}{2}\right)\ln\left(\frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|}\right) \tag{6-43}$$

has approximately a chi-square distribution with $p(g-1)$ d.f. Consequently, for $\Sigma n_\ell = n$ large, we reject H_0 at significance level α if

$$-\left(n - 1 - \frac{(p+g)}{2}\right)\ln\left(\frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|}\right) > \chi^2_{p(g-1)}(\alpha) \tag{6-44}$$

where $\chi^2_{p(g-1)}(\alpha)$ is the upper (100α)th percentile of a chi-square distribution with $p(g-1)$ d.f.

or $\left(\frac{\Sigma n_\ell - p - 2}{p}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \leq F_{2(p),2(N)}(0.01)$

Table 6.3 Distribution of Wilks' Lambda, $\Lambda^* = |\mathbf{W}|/|\mathbf{B} + \mathbf{W}|$

No. of variables	No. of groups	Sampling distribution for multivariate normal data
$p = 1$	$g \geq 2$	$\left(\frac{\Sigma n_\ell - g}{g-1}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{g-1,\Sigma n_\ell - g}$
$p = 2$	$g \geq 2$	$\left(\frac{\Sigma n_\ell - g - 1}{g-1}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(g-1),2(\Sigma n_\ell - g - 1)}$
$p \geq 1$	$g = 2$	$\left(\frac{\Sigma n_\ell - p - 1}{p}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{p,\Sigma n_\ell - p - 1}$
$p \geq 1$	$g = 3$	$\left(\frac{\Sigma n_\ell - p - 2}{p}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2p,2(\Sigma n_\ell - p - 2)}$

Simultaneous Confidence Intervals for Treatment Effects

- When the hypothesis of equal treatment effects is rejected, those effects that led to the rejection of the hypothesis are of interest.

It remains to apportion the error rate over the numerous confidence statements. Relation (5-28) still applies. There are p variables and $g(g-1)/2$ pairwise differences, so each two-sample t-interval will employ the critical value $t_{n-g}(\alpha/2m)$, where

$$m = pg(g-1)/2 \tag{6-46}$$

is the number of simultaneous confidence statements.

Result 6.5. Let $n = \sum_{k=1}^{g} n_k$. For the model in (6-38), with confidence at least $(1-\alpha)$,

$$\tau_{ki} - \tau_{\ell i} \quad \text{belongs to} \quad \bar{x}_{ki} - \bar{x}_{\ell i} \pm t_{n-g}\left(\frac{\alpha}{pg(g-1)}\right)\sqrt{\frac{w_{ii}}{n-g}\left(\frac{1}{n_k} + \frac{1}{n_\ell}\right)}$$

for all components $i = 1, \ldots, p$ and all differences $\ell < k = 1, \ldots, g$. Here w_{ii} is the ith diagonal element of $\mathbf{W}$.

Testing for Equality of Covariance Matrices

- One of the assumptions made when comparing two or more multivariate mean vectors is that the covariance matrices of the potentially different populations are the same.

Box's M-test:

- testing for equal covariance

With g populations, the null hypothesis is

$$H_0: \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_g = \boldsymbol{\Sigma} \tag{6-47}$$

where $\boldsymbol{\Sigma}_\ell$ is the covariance matrix for the ℓth population, $\ell = 1, 2, \ldots, g$, and $\boldsymbol{\Sigma}$ is the presumed common covariance matrix. The alternative hypothesis is that at least two of the covariance matrices are not equal.

- Likelihood Ratio Statistic for testing equal variance

$$\Lambda = \prod_{\ell}\left(\frac{|S_\ell|}{|S_{pooled}|}\right)^{\frac{n_\ell - 1}{2}}, \quad S_{pooled} = \frac{1}{\Sigma(n_\ell - 1)}\{(n_1-1)S_1 + (n_2-1)S_2 + \cdots + (n_g-1)S_g\}$$

Box's test is based on his χ^2 approximation to the sampling distribution of $-2\ln\Lambda$ (see Result 5.2). Setting $-2\ln\Lambda = M$ (Box's M statistic) gives

$$M = \left[\sum_\ell (n_\ell - 1)\right]\ln|\mathbf{S}_{\text{pooled}}| - \sum_\ell [(n_\ell - 1)\ln|\mathbf{S}_\ell|] \tag{6-50}$$

- note that the determinant of the pooled covariance matrix $|S_{pooled}|$ will lie somewhere near the "middle" of the determinants $|S_\ell|$'s.

Box's Test for Equality of Covariance Matrices

Set

$$u = \left[\sum_{\ell} \frac{1}{(n_\ell - 1)} - \frac{1}{\sum_{\ell}(n_\ell - 1)} \right] \left[\frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \right] \tag{6-51}$$

where p is the number of variables and g is the number of groups. Then

$$C = (1-u)M = (1-u)\left\{ \left[\sum_{\ell}(n_\ell - 1) \right] \ln |\mathbf{S}_{\text{pooled}}| - \sum_{\ell} [(n_\ell - 1) \ln |\mathbf{S}_\ell|] \right\} \tag{6-52}$$

has an approximate χ^2 distribution with

$$\nu = g\frac{1}{2}p(p+1) - \frac{1}{2}p(p+1) = \frac{1}{2}p(p+1)(g-1) \tag{6-53}$$

degrees of freedom. At significance level α, reject H_0 if $C > \chi^2_{p(p+1)(g-1)/2}(\alpha)$.

Box's χ^2 approximation works well if each n_ℓ exceeds 20 and if p and g do not exceed 5. In situations where these conditions do not hold, Box ([7], [8]) has provided a more precise F approximation to the sampling distribution of M.

Two-Way Multivariate Analysis of Variance

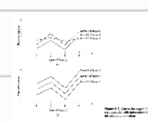
Univariate Two-Way Fixed-Effects Model with Interaction

- Suppose there are g levels of factor 1 and b levels of factor 2, and n independent observations can be observed at each of the gb combinations of levels.

$$\underbrace{X_{\ell k r}}_{\text{response}} = \underbrace{\mu}_{\text{overall level}} + \underbrace{\tau_\ell}_{\text{effect of factor 1}} + \underbrace{\beta_k}_{\text{effect of factor 2}} + \underbrace{\gamma_{\ell k}}_{\text{factor 1\&2 interaction}} + e_{\ell k r}, \quad \ell = 1, 2, \ldots, g \quad k = 1, 2, \ldots, b \quad r = 1, 2, \ldots, n$$

$$* \sum_{\ell=1}^{g} \tau_\ell = \sum_{k=1}^{b} \beta_k = \sum_{\ell=1}^{g} \gamma_{\ell k} = \sum_{k=1}^{b} \gamma_{\ell k} = 0$$

$$\to X_{\ell k r} = \bar{x} + (\bar{x}_{\ell \cdot} - \bar{x}) + (\bar{x}_{\cdot k} - \bar{x}) + (\underbrace{\bar{x}_{\ell k}}_{\text{Average for the } \ell^{th} \text{ level of factor 1 and the } k^{th} \text{ level of factor 2}} - \bar{x}_{\ell} - \bar{x}_{\cdot k} + \bar{x}) + (X_{\ell k r} - \bar{x}_{\ell k})$$



factor 1 *and* the kth level of factor 2. Squaring and summing the deviations $(x_{\ell k r} - \bar{x})$ gives

$$\begin{aligned} \sum_{\ell=1}^{g} \sum_{k=1}^{b} \sum_{r=1}^{n} (x_{\ell k r} - \bar{x})^2 &= \sum_{\ell=1}^{g} bn(\bar{x}_{\ell \cdot} - \bar{x})^2 + \sum_{k=1}^{b} gn(\bar{x}_{\cdot k} - \bar{x})^2 \\ &+ \sum_{\ell=1}^{g} \sum_{k=1}^{b} n(\bar{x}_{\ell k} - \bar{x}_{\ell \cdot} - \bar{x}_{\cdot k} + \bar{x})^2 \\ &+ \sum_{\ell=1}^{g} \sum_{k=1}^{b} \sum_{r=1}^{n} (x_{\ell k r} - \bar{x}_{\ell k})^2 \end{aligned} \tag{6-57}$$

or

$$SS_{\text{cor}} = SS_{\text{fac1}} + SS_{\text{fac2}} + SS_{\text{int}} + SS_{\text{res}}$$

The corresponding degrees of freedom associated with the sums of squares in the breakup in (6-57) are

$$gbn - 1 = (g-1) + (b-1) + (g-1)(b-1) + gb(n-1) \tag{6-58}$$

ANOVA Table for Comparing Effects of Two Factors and Their Interaction

Source of variation	Sum of squares (SS)	Degrees of freedom (d.f.)
Factor 1	$SS_{fac1} = \sum_{\ell=1}^{g} bn(\bar{x}_{\ell\cdot} - \bar{x})^2$	$g - 1$
Factor 2	$SS_{fac2} = \sum_{k=1}^{b} gn(\bar{x}_{\cdot k} - \bar{x})^2$	$b - 1$
Interaction	$SS_{int} = \sum_{\ell=1}^{g}\sum_{k=1}^{b} n(\bar{x}_{\ell k} - \bar{x}_{\ell\cdot} - \bar{x}_{\cdot k} + \bar{x})^2$	$(g-1)(b-1)$
Residual (Error)	$SS_{res} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(x_{\ell kr} - \bar{x}_{\ell k})^2$	$gb(n-1)$
Total (corrected)	$SS_{cor} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(x_{\ell kr} - \bar{x})^2$	$gbn - 1$

The F-ratios of the mean squares, $SS_{fac1}/(g-1)$, $SS_{fac2}/(b-1)$, and $SS_{int}/(g-1)(b-1)$ to the mean square, $SS_{res}/(gb(n-1))$ can be used to test for the effects of factor 1, factor 2, and factor 1–factor 2 interaction, respectively. (See

Multivariate Two-Way Fixed-Effects Model with Interaction

Again, the generalization from the univariate to the multivariate analysis consists simply of replacing a scalar such as $(\bar{x}_{\ell\cdot} - \bar{x})^2$ with the corresponding matrix $(\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}})(\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}})'$.

The MANOVA table is the following:

MANOVA Table for Comparing Factors and Their Interaction

Source of variation	Matrix of sum of squares and cross products (SSP)	Degrees of freedom (d.f.)
Factor 1	$SSP_{fac1} = \sum_{\ell=1}^{g} bn(\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}})(\bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}})'$	$g - 1$
Factor 2	$SSP_{fac2} = \sum_{k=1}^{b} gn(\bar{\mathbf{x}}_{\cdot k} - \bar{\mathbf{x}})(\bar{\mathbf{x}}_{\cdot k} - \bar{\mathbf{x}})'$	$b - 1$
Interaction	$SSP_{int} = \sum_{\ell=1}^{g}\sum_{k=1}^{b} n(\bar{\mathbf{x}}_{\ell k} - \bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}}_{\cdot k} + \bar{\mathbf{x}})(\bar{\mathbf{x}}_{\ell k} - \bar{\mathbf{x}}_{\ell\cdot} - \bar{\mathbf{x}}_{\cdot k} + \bar{\mathbf{x}})'$	$(g-1)(b-1)$
Residual (Error)	$SSP_{res} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(\mathbf{x}_{\ell kr} - \bar{\mathbf{x}}_{\ell k})(\mathbf{x}_{\ell kr} - \bar{\mathbf{x}}_{\ell k})'$	$gb(n-1)$
Total (corrected)	$SSP_{cor} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(\mathbf{x}_{\ell kr} - \bar{\mathbf{x}})(\mathbf{x}_{\ell kr} - \bar{\mathbf{x}})'$	$gbn - 1$

여기서 의미하는 n은 전체 N이 아닌 Factors들이 같은 각 그룹의 sample 수

- A test is conducted by, $H_0: \gamma_{11} = \gamma_{12} = \cdots = \gamma_{gb} = 0$ (No interaction effects), and it is (the likelihood ratio test) rejected if Wilk's lambda, $\Lambda^* = \dfrac{|SSP_{res}|}{|SSP_{int} + SSP_{res}|}$, applied to

$$-\left[gb(n-1) - \frac{p+1-(g-1)(b-1)}{2}\right]\ln\Lambda^* > \chi^2_{(g-1)(b-1)p}(\alpha)$$

, is greater than χ^2 statistic

- Ordinarily, the test for interaction is carried out before the tests for main factor effects. If interaction effects exist, p univariate two-way analyses of variance are often conducted to see whether the interaction appears in some responses but not others, and those responses without interaction may be interpreted in terms of additive factor 1 and 2 effects.

Testing for factor 1 and factor 2

In the multivariate model, we test for factor 1 and factor 2 main effects as follows. First, consider the hypotheses H_0: $\boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \mathbf{0}$ and H_1: at least one $\boldsymbol{\tau}_\ell \neq \mathbf{0}$. These hypotheses specify *no* factor 1 effects and *some* factor 1 effects, respectively. Let

$$\Lambda^* = \frac{|\mathrm{SSP}_{\mathrm{res}}|}{|\mathrm{SSP}_{\mathrm{fac1}} + \mathrm{SSP}_{\mathrm{res}}|} \tag{6-66}$$

so that small values of Λ^* are consistent with H_1. Using Bartlett's correction, the likelihood ratio test is as follows:

Reject H_0: $\boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \mathbf{0}$ (no factor 1 effects) at level α if

$$-\left[gb(n-1) - \frac{p+1-(g-1)}{2}\right]\ln \Lambda^* > \chi^2_{(g-1)p}(\alpha) \tag{6-67}$$

where Λ^* is given by (6-66) and $\chi^2_{(g-1)p}(\alpha)$ is the upper (100α)th percentile of a chi-square distribution with $(g-1)p$ d.f.

In a similar manner, factor 2 effects are tested by considering H_0: $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_b = \mathbf{0}$ and H_1: at least one $\boldsymbol{\beta}_k \neq \mathbf{0}$. Small values of

$$\Lambda^* = \frac{|\mathrm{SSP}_{\mathrm{res}}|}{|\mathrm{SSP}_{\mathrm{fac2}} + \mathrm{SSP}_{\mathrm{res}}|} \tag{6-68}$$

are consistent with H_1. Once again, for large samples and using Bartlett's correction: Reject H_0: $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_b = \mathbf{0}$ (no factor 2 effects) at level α if

$$-\left[gb(n-1) - \frac{p+1-(b-1)}{2}\right]\ln \Lambda^* > \chi^2_{(b-1)p}(\alpha) \tag{6-69}$$

where Λ^* is given by (6-68) and $\chi^2_{(b-1)p}(\alpha)$ is the upper (100α)th percentile of a chi-square distribution with $(b-1)p$ degrees of freedom.

Simultaneous Confidence Intervals for $\tau_{\ell i} - \tau_{mi}$

The $100(1-\alpha)\%$ simultaneous confidence intervals for $\tau_{\ell i} - \tau_{mi}$ are

$$\tau_{\ell i} - \tau_{mi} \quad \text{belongs to} \quad (\bar{x}_{\ell \cdot i} - \bar{x}_{m \cdot i}) \pm t_\nu\left(\frac{\alpha}{pg(g-1)}\right)\sqrt{\frac{E_{ii}}{\nu}\frac{2}{bn}} \tag{6-70}$$

where $\nu = gb(n-1)$, E_{ii} is the ith diagonal element of $\mathbf{E} = \mathrm{SSP}_{\mathrm{res}}$, and $\bar{x}_{\ell \cdot i} - \bar{x}_{m \cdot i}$ is the ith component of $\bar{\mathbf{x}}_{\ell \cdot} - \bar{\mathbf{x}}_{m \cdot}$.

Similarly, the $100(1-\alpha)$ percent simultaneous confidence intervals for $\beta_{ki} - \beta_{qi}$ are

$$\beta_{ki} - \beta_{qi} \quad \text{belongs to} \quad (\bar{x}_{\cdot ki} - \bar{x}_{\cdot qi}) \pm t_\nu\left(\frac{\alpha}{pb(b-1)}\right)\sqrt{\frac{E_{ii}}{\nu}\frac{2}{gn}} \tag{6-71}$$

where ν and E_{ii} are as just defined and $\bar{x}_{\cdot ki} - \bar{x}_{\cdot qi}$ is the ith component of $\bar{\mathbf{x}}_{\cdot k} - \bar{\mathbf{x}}_{\cdot q}$.