A very short proof of the Multivariate Chebyshev's Inequality. Applications to order statistics and data sets.

J. Navarro[1], Universidad de Murcia (Spain).

[1] Supported by Ministerio de Economía y Competitividad under Grant MTM2012-34023-FEDER.

Notation

- $\mathbf{X} = (X_1, \ldots, X_k)'$ a random vector.
- $\mu = E(\mathbf{X}) = (\mu_1, \ldots, \mu_k)'$ mean vector.
- $V = Cov(\mathbf{X}) = E((\mathbf{X} - \mu)(\mathbf{X} - \mu)')$ covariance matrix.
- $\mathbf{x} = (x_1, \ldots, x_k)' \in \mathbb{R}^k$.
- Mahalanobis distance from $\mathbf{x}$ to μ:

$$\Delta_V(\mathbf{x}, \mu) = \sqrt{(\mathbf{x} - \mu)' V^{-1} (\mathbf{x} - \mu)}.$$

Notation

- $\mathbf{X} = (X_1, \ldots, X_k)'$ a random vector.
- $\boldsymbol{\mu} = E(\mathbf{X}) = (\mu_1, \ldots, \mu_k)'$ mean vector.
- $V = Cov(\mathbf{X}) = E((\mathbf{X} - \mu)(\mathbf{X} - \mu)')$ covariance matrix.
- $\mathbf{x} = (x_1, \ldots, x_k)' \in \mathbb{R}^k$.
- Mahalanobis distance from $\mathbf{x}$ to $\boldsymbol{\mu}$:

$$\Delta_V(\mathbf{x}, \boldsymbol{\mu}) = \sqrt{(\mathbf{x} - \boldsymbol{\mu})' V^{-1} (\mathbf{x} - \boldsymbol{\mu})}.$$

Notation

- $\mathbf{X} = (X_1, \ldots, X_k)'$ a random vector.
- $\boldsymbol{\mu} = E(\mathbf{X}) = (\mu_1, \ldots, \mu_k)'$ mean vector.
- $V = Cov(\mathbf{X}) = E((\mathbf{X} - \mu)(\mathbf{X} - \mu)')$ covariance matrix.
- $\mathbf{x} = (x_1, \ldots, x_k)' \in \mathbb{R}^k$.
- Mahalanobis distance from $\mathbf{x}$ to μ:

$$\Delta_V(\mathbf{x}, \mu) = \sqrt{(\mathbf{x} - \mu)' V^{-1} (\mathbf{x} - \mu)}.$$

J. Navarro, E-mail: jorgenav@um.es

Notation

- $\mathbf{X} = (X_1, \ldots, X_k)'$ a random vector.
- $\boldsymbol{\mu} = E(\mathbf{X}) = (\mu_1, \ldots, \mu_k)'$ mean vector.
- $V = Cov(\mathbf{X}) = E((\mathbf{X} - \mu)(\mathbf{X} - \mu)')$ covariance matrix.
- $\mathbf{x} = (x_1, \ldots, x_k)' \in \mathbb{R}^k$.
- Mahalanobis distance from $\mathbf{x}$ to μ:

$$\Delta_V(\mathbf{x}, \mu) = \sqrt{(\mathbf{x} - \mu)' V^{-1} (\mathbf{x} - \mu)}.$$

Notation

- $\mathbf{X} = (X_1, \ldots, X_k)'$ a random vector.
- $\boldsymbol{\mu} = E(\mathbf{X}) = (\mu_1, \ldots, \mu_k)'$ mean vector.
- $V = Cov(\mathbf{X}) = E((\mathbf{X} - \mu)(\mathbf{X} - \mu)')$ covariance matrix.
- $\mathbf{x} = (x_1, \ldots, x_k)' \in \mathbb{R}^k$.
- Mahalanobis distance from $\mathbf{x}$ to $\boldsymbol{\mu}$:

$$\Delta_V(\mathbf{x}, \boldsymbol{\mu}) = \sqrt{(\mathbf{x} - \boldsymbol{\mu})' V^{-1} (\mathbf{x} - \boldsymbol{\mu})}.$$

The (univariate) Markov's inequality.

- If Z is a non-negative random variable with finite mean $E(Z)$ and $\varepsilon > 0$, then

$$\varepsilon \Pr(Z \geq \varepsilon) = \varepsilon \int_{[\varepsilon,\infty)} dF_Z(x) \leq \int_{[\varepsilon,\infty)} x dF_Z(x) \leq \int_{[0,\infty)} x dF_Z(x) = E(Z)$$

  where $F_Z(x) = \Pr(Z \leq x)$.
- It can be stated as

$$\Pr(Z \geq \varepsilon) \leq \frac{E(Z)}{\varepsilon}. \quad (1)$$

J. Navarro, E-mail: jorgenav@um.es

The (univariate) Markov's inequality.

- If Z is a non-negative random variable with finite mean $E(Z)$ and $\varepsilon > 0$, then

$$\varepsilon \Pr(Z \geq \varepsilon) = \varepsilon \int_{[\varepsilon,\infty)} dF_Z(x) \leq \int_{[\varepsilon,\infty)} x dF_Z(x) \leq \int_{[0,\infty)} x dF_Z(x) = E(Z)$$

where $F_Z(x) = \Pr(Z \leq x)$.

- It can be stated as

$$\Pr(Z \geq \varepsilon) \leq \frac{E(Z)}{\varepsilon}. \qquad (1)$$

The univariate Chebyshev's inequality.

- If X is a random variable with finite mean $\mu = E(X)$ and variance $\sigma^2 = Var(X) > 0$, then by taking $Z = (X - \mu)^2/\sigma^2$ in (1), we get

$$\Pr\left(\frac{(X-\mu)^2}{\sigma^2} \geq \varepsilon\right) \leq \frac{1}{\varepsilon} \qquad (2)$$

for all $\varepsilon > 0$.

- It can also be written as

$$\Pr((X-\mu)^2 < \varepsilon\sigma^2) \geq 1 - \frac{1}{\varepsilon}$$

or as

$$\Pr(|X-\mu| < r) \leq 1 - \frac{\sigma^2}{r^2}$$

for all $r > 0$.

The univariate Chebyshev's inequality.

- If X is a random variable with finite mean $\mu = E(X)$ and variance $\sigma^2 = Var(X) > 0$, then by taking $Z = (X - \mu)^2/\sigma^2$ in (1), we get

$$\Pr\left(\frac{(X-\mu)^2}{\sigma^2} \geq \varepsilon\right) \leq \frac{1}{\varepsilon} \tag{2}$$

  for all $\varepsilon > 0$.

- It can also be written as

$$\Pr((X-\mu)^2 < \varepsilon\sigma^2) \geq 1 - \frac{1}{\varepsilon}$$

  or as

$$\Pr(|X-\mu| < r) \leq 1 - \frac{\sigma^2}{r^2}$$

  for all $r > 0$.

The multivariate Chebyshev's inequality (MCI).

- If $\mathbf{X}$ is a random vector with finite mean $\boldsymbol{\mu} = E(\mathbf{X})'$ and positive definite covariance matrix $V = Cov(\mathbf{X})$.
- Then
$$\Pr((\mathbf{X}-\mu)'V^{-1}(\mathbf{X}-\mu) \geq \varepsilon) \leq \frac{k}{\varepsilon} \tag{3}$$
for all $\varepsilon > 0$.
- Chen, X. (2011). A new generalization of Chebyshev inequality for random vectors. arXiv:0707.0805v2.

The multivariate Chebyshev's inequality (MCI).

- If $\mathbf{X}$ is a random vector with finite mean $\boldsymbol{\mu} = E(\mathbf{X})'$ and positive definite covariance matrix $V = Cov(\mathbf{X})$.
- Then

$$\Pr((\mathbf{X}-\boldsymbol{\mu})'V^{-1}(\mathbf{X}-\boldsymbol{\mu}) \geq \varepsilon) \leq \frac{k}{\varepsilon} \tag{3}$$

for all $\varepsilon > 0$.

- Chen, X. (2011). A new generalization of Chebyshev inequality for random vectors. arXiv:0707.0805v2.

The multivariate Chebyshev's inequality (MCI).

- If $\mathbf{X}$ is a random vector with finite mean $\boldsymbol{\mu} = E(\mathbf{X})'$ and positive definite covariance matrix $V = Cov(\mathbf{X})$.
- Then

$$\Pr((\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}) \geq \varepsilon) \leq \frac{k}{\varepsilon} \tag{3}$$

  for all $\varepsilon > 0$.
- Chen, X. (2011). A new generalization of Chebyshev inequality for random vectors. arXiv:0707.0805v2.

The multivariate Chebyshev's inequality.

- The inequality in (3) can also be written as

$$\Pr((\mathbf{X}-\boldsymbol{\mu})'V^{-1}(\mathbf{X}-\boldsymbol{\mu})<\varepsilon)\geq 1-\frac{k}{\varepsilon} \tag{4}$$

for all $\varepsilon > 0$.

- This inequality says that the ellipsoid

$$E_\varepsilon = \{\mathbf{x}\in\mathbb{R}^k : (\mathbf{x}-\boldsymbol{\mu})'V^{-1}(\mathbf{x}-\boldsymbol{\mu})<\varepsilon\} \tag{5}$$

contains at least the $100(1-k/\varepsilon)\%$ of the population.

- The inequality can also be written as

$$\Pr(\Delta_V(\mathbf{X},\boldsymbol{\mu})<r)\geq 1-\frac{k}{r^2}. \tag{6}$$

- Hence (6) gives a lower bound for the percentage of points from $\mathbf{X}$ in spheres "around" the mean $\boldsymbol{\mu}$ in the Mahalanobis distance based on V.

The multivariate Chebyshev's inequality.

- The inequality in (3) can also be written as

$$\Pr((\mathbf{X}-\boldsymbol{\mu})'V^{-1}(\mathbf{X}-\boldsymbol{\mu})<\varepsilon)\geq 1-\frac{k}{\varepsilon} \quad (4)$$

for all $\varepsilon > 0$.

- This inequality says that the ellipsoid

$$E_\varepsilon = \{\mathbf{x}\in\mathbb{R}^k : (\mathbf{x}-\boldsymbol{\mu})'V^{-1}(\mathbf{x}-\boldsymbol{\mu})<\varepsilon\} \quad (5)$$

contains at least the $100(1-k/\varepsilon)\%$ of the population.

- The inequality can also be written as

$$\Pr(\Delta_V(\mathbf{X},\boldsymbol{\mu})<r)\geq 1-\frac{k}{r^2}. \quad (6)$$

- Hence (6) gives a lower bound for the percentage of points from $\mathbf{X}$ in spheres "around" the mean $\boldsymbol{\mu}$ in the Mahalanobis distance based on V.

The multivariate Chebyshev's inequality.

- The inequality in (3) can also be written as

$$\Pr((\mathbf{X}-\boldsymbol{\mu})'V^{-1}(\mathbf{X}-\boldsymbol{\mu})<\varepsilon)\geq 1-\frac{k}{\varepsilon} \tag{4}$$

  for all $\varepsilon > 0$.
- This inequality says that the ellipsoid

$$E_\varepsilon=\{\mathbf{x}\in\mathbb{R}^k:(\mathbf{x}-\boldsymbol{\mu})'V^{-1}(\mathbf{x}-\boldsymbol{\mu})<\varepsilon\} \tag{5}$$

  contains at least the $100(1-k/\varepsilon)\%$ of the population.
- The inequality can also be written as

$$\Pr(\Delta_V(\mathbf{X},\boldsymbol{\mu})<r)\geq 1-\frac{k}{r^2}. \tag{6}$$

- Hence (6) gives a lower bound for the percentage of points from $\mathbf{X}$ in spheres "around" the mean μ in the Mahalanobis distance based on V.

The multivariate Chebyshev's inequality.

- The inequality in (3) can also be written as

$$\Pr((\mathbf{X}-\boldsymbol{\mu})'V^{-1}(\mathbf{X}-\boldsymbol{\mu})<\varepsilon)\geq 1-\frac{k}{\varepsilon} \tag{4}$$

  for all $\varepsilon > 0$.
- This inequality says that the ellipsoid

$$E_\varepsilon=\{\mathbf{x}\in\mathbb{R}^k:(\mathbf{x}-\boldsymbol{\mu})'V^{-1}(\mathbf{x}-\boldsymbol{\mu})<\varepsilon\} \tag{5}$$

  contains at least the $100(1-k/\varepsilon)\%$ of the population.
- The inequality can also be written as

$$\Pr(\Delta_V(\mathbf{X},\boldsymbol{\mu})<r)\geq 1-\frac{k}{r^2}. \tag{6}$$

- Hence (6) gives a lower bound for the percentage of points from $\mathbf{X}$ in spheres "around" the mean $\boldsymbol{\mu}$ in the Mahalanobis distance based on V.

A very short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}).$$

- As V is positive definite, then $Z \geq 0$.
- Moreover, there exist symmetric matrices $V^{1/2}$ and $V^{-1/2}$ such that $V^{1/2}V^{1/2} = V$, $V^{-1/2}V^{-1/2} = V^{-1}$ and $V^{1/2}V^{-1/2} = V^{-1/2}V^{1/2} = I_k$, where I_k is the identity matrix of dimension k.
- Therefore

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1/2} V^{-1/2} (\mathbf{X} - \boldsymbol{\mu}) = \mathbf{Y}'\mathbf{Y},$$

where $\mathbf{Y} = (Y_1, \dots, Y_k)' = V^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$.
- Hence $E(\mathbf{Y}) = \mathbf{0}_k$ and

$$Cov(\mathbf{Y}) = V^{-1/2} Cov(\mathbf{X}) V^{-1/2} = V^{-1/2} V V^{-1/2} = I_k.$$

A very short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}).$$

- As V is positive definite, then $Z \geq 0$.
- Moreover, there exist symmetric matrices $V^{1/2}$ and $V^{-1/2}$ such that $V^{1/2}V^{1/2} = V$, $V^{-1/2}V^{-1/2} = V^{-1}$ and $V^{1/2}V^{-1/2} = V^{-1/2}V^{1/2} = I_k$, where I_k is the identity matrix of dimension k.
- Therefore

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1/2} V^{-1/2} (\mathbf{X} - \boldsymbol{\mu}) = \mathbf{Y}'\mathbf{Y},$$

where $\mathbf{Y} = (Y_1, \ldots, Y_k)' = V^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$.
- Hence $E(\mathbf{Y}) = \mathbf{0}_k$ and

$$Cov(\mathbf{Y}) = V^{-1/2} Cov(\mathbf{X}) V^{-1/2} = V^{-1/2} V V^{-1/2} = I_k.$$

A very short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}).$$

- As V is positive definite, then $Z \geq 0$.
- Moreover, there exist symmetric matrices $V^{1/2}$ and $V^{-1/2}$ such that $V^{1/2}V^{1/2} = V$, $V^{-1/2}V^{-1/2} = V^{-1}$ and $V^{1/2}V^{-1/2} = V^{-1/2}V^{1/2} = I_k$, where I_k is the identity matrix of dimension k.
- Therefore

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1/2} V^{-1/2} (\mathbf{X} - \boldsymbol{\mu}) = \mathbf{Y}'\mathbf{Y},$$

where $\mathbf{Y} = (Y_1, \ldots, Y_k)' = V^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$.
- Hence $E(\mathbf{Y}) = \mathbf{0}_k$ and

$$Cov(\mathbf{Y}) = V^{-1/2} Cov(\mathbf{X}) V^{-1/2} = V^{-1/2} V V^{-1/2} = I_k.$$

A very short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}).$$

- As V is positive definite, then $Z \geq 0$.
- Moreover, there exist symmetric matrices $V^{1/2}$ and $V^{-1/2}$ such that $V^{1/2}V^{1/2} = V$, $V^{-1/2}V^{-1/2} = V^{-1}$ and $V^{1/2}V^{-1/2} = V^{-1/2}V^{1/2} = I_k$, where I_k is the identity matrix of dimension k.
- Therefore

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1/2} V^{-1/2} (\mathbf{X} - \boldsymbol{\mu}) = \mathbf{Y}'\mathbf{Y},$$

where $\mathbf{Y} = (Y_1, \ldots, Y_k)' = V^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$.
- Hence $E(\mathbf{Y}) = \mathbf{0}_k$ and

$$Cov(\mathbf{Y}) = V^{-1/2} Cov(\mathbf{X}) V^{-1/2} = V^{-1/2} V V^{-1/2} = I_k.$$

A very short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}).$$

- As V is positive definite, then $Z \geq 0$.
- Moreover, there exist symmetric matrices $V^{1/2}$ and $V^{-1/2}$ such that $V^{1/2}V^{1/2} = V$, $V^{-1/2}V^{-1/2} = V^{-1}$ and $V^{1/2}V^{-1/2} = V^{-1/2}V^{1/2} = I_k$, where I_k is the identity matrix of dimension k.
- Therefore

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1/2} V^{-1/2} (\mathbf{X} - \boldsymbol{\mu}) = \mathbf{Y}'\mathbf{Y},$$

where $\mathbf{Y} = (Y_1, \ldots, Y_k)' = V^{-1/2}(\mathbf{X} - \boldsymbol{\mu})$.
- Hence $E(\mathbf{Y}) = \mathbf{0}_k$ and

$$Cov(\mathbf{Y}) = V^{-1/2} Cov(\mathbf{X}) V^{-1/2} = V^{-1/2} V V^{-1/2} = I_k.$$

A very short proof.

- Therefore $E(Y_i) = 0$, $Var(Y_i) = 1$ and

$$E(Z) = E(\mathbf{Y}'\mathbf{Y}) = E\left(\sum_{i=1}^{k} Y_i^2\right) = \sum_{i=1}^{k} E(Y_i^2) = \sum_{i=1}^{k} Var(Y_i) = k.$$

- Hence, from Markov's inequality (1), we get

$$\Pr(Z \geq \varepsilon) = \Pr((\mathbf{X}-\mu)' V^{-1}(\mathbf{X}-\mu) \geq \varepsilon) \leq \frac{E(Z)}{\varepsilon} = \frac{k}{\varepsilon}$$

and therefore (3) holds for all $\varepsilon > 0$.

A very short proof.

- Therefore $E(Y_i) = 0$, $Var(Y_i) = 1$ and

$$E(Z) = E(\mathbf{Y}'\mathbf{Y}) = E\left(\sum_{i=1}^{k} Y_i^2\right) = \sum_{i=1}^{k} E(Y_i^2) = \sum_{i=1}^{k} Var(Y_i) = k.$$

- Hence, from Markov's inequality (1), we get

$$\Pr(Z \geq \varepsilon) = \Pr((\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}) \geq \varepsilon) \leq \frac{E(Z)}{\varepsilon} = \frac{k}{\varepsilon}$$

and therefore (3) holds for all $\varepsilon > 0$.

- That's all, thank you for your attention!!
- It's a joke, let's see something more (if you want).

- That's all, thank you for your attention!!
- It's a joke, let's see something more (if you want).

Another short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}) \geq 0.$$

- As V is positive definite and symmetric, there exists an ortogonal matrix T such that $TT' = T'T = I_k$ and $T'VT = D$ and $D = diag(\lambda_1, \ldots, \lambda_k)$ is the diagonal matrix with the ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_k > 0$.
- Then $V = TDT'$ and $V^{-1} = TD^{-1}T'$.
- Therefore

$$\begin{aligned} Z &= (\mathbf{X} - \mu)' TD^{-1} T' (\mathbf{X} - \mu) \\ &= [D^{-1/2} T' (\mathbf{X} - \mu)]' [D^{-1/2} T' (\mathbf{X} - \mu)] \\ &= \mathbf{Z}'\mathbf{Z}, \end{aligned}$$

where $\mathbf{Z} = (Z_1, \ldots, Z_n)' = D^{-1/2} T' (\mathbf{X} - \boldsymbol{\mu})$ and $D^{-1/2} = diag(\lambda_1^{-1/2}, \ldots, \lambda_k^{-1/2})$.

J. Navarro, E-mail: jorgenav@um.es

Another short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}) \geq 0.$$

- As V is positive definite and symmetric, there exists an ortogonal matrix T such that $TT' = T'T = I_k$ and $T'VT = D$ and $D = diag(\lambda_1, \ldots, \lambda_k)$ is the diagonal matrix with the ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_k > 0$.
- Then $V = TDT'$ and $V^{-1} = TD^{-1}T'$.
- Therefore

$$\begin{aligned} Z &= (\mathbf{X} - \mu)' T D^{-1} T' (\mathbf{X} - \mu) \\ &= [D^{-1/2} T' (\mathbf{X} - \mu)]' [D^{-1/2} T' (\mathbf{X} - \mu)] \\ &= \mathbf{Z}'\mathbf{Z}, \end{aligned}$$

where $\mathbf{Z} = (Z_1, \ldots, Z_n)' = D^{-1/2} T' (\mathbf{X} - \mu)$ and $D^{-1/2} = diag(\lambda_1^{-1/2}, \ldots, \lambda_k^{-1/2})$.

Another short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}) \geq 0.$$

- As V is positive definite and symmetric, there exists an ortogonal matrix T such that $TT' = T'T = I_k$ and $T'VT = D$ and $D = diag(\lambda_1, \dots, \lambda_k)$ is the diagonal matrix with the ordered eigenvalues $\lambda_1 \geq \dots \geq \lambda_k > 0$.
- Then $V = TDT'$ and $V^{-1} = TD^{-1}T'$.
- Therefore

$$\begin{aligned} Z &= (\mathbf{X} - \mu)' T D^{-1} T' (\mathbf{X} - \mu) \\ &= [D^{-1/2} T' (\mathbf{X} - \mu)]' [D^{-1/2} T' (\mathbf{X} - \mu)] \\ &= \mathbf{Z}'\mathbf{Z}, \end{aligned}$$

where $\mathbf{Z} = (Z_1, \dots, Z_n)' = D^{-1/2} T' (\mathbf{X} - \mu)$ and $D^{-1/2} = diag(\lambda_1^{-1/2}, \dots, \lambda_k^{-1/2})$.

Another short proof.

- Let us consider the random variable

$$Z = (\mathbf{X} - \boldsymbol{\mu})' V^{-1} (\mathbf{X} - \boldsymbol{\mu}) \geq 0.$$

- As V is positive definite and symmetric, there exists an ortogonal matrix T such that $TT' = T'T = I_k$ and $T'VT = D$ and $D = diag(\lambda_1, \ldots, \lambda_k)$ is the diagonal matrix with the ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_k > 0$.
- Then $V = TDT'$ and $V^{-1} = TD^{-1}T'$.
- Therefore

$$\begin{aligned} Z &= (\mathbf{X} - \mu)' TD^{-1} T' (\mathbf{X} - \boldsymbol{\mu}) \\ &= [D^{-1/2} T' (\mathbf{X} - \boldsymbol{\mu})]' [D^{-1/2} T' (\mathbf{X} - \boldsymbol{\mu})] \\ &= \mathbf{Z}'\mathbf{Z}, \end{aligned}$$

where $\mathbf{Z} = (Z_1, \ldots, Z_n)' = D^{-1/2} T' (\mathbf{X} - \boldsymbol{\mu})$ and $D^{-1/2} = diag(\lambda_1^{-1/2}, \ldots, \lambda_k^{-1/2})$.

Another short proof.

- The random vector $\mathbf{Z}$ satisfies $E(\mathbf{Z}) = \mathbf{0}_k$ and

$$Cov(\mathbf{Z}) = Cov(D^{-1/2}T'(\mathbf{X}-\mu)) = D^{-1/2}T'VTD^{-1/2} = D^{-1/2}DD^{-1/2} = I_k.$$

- Therefore $E(Z_i) = 0$, $Var(Z_i) = 1$ and

$$E(Z) = E(\mathbf{Z}'\mathbf{Z}) = E\left(\sum_{i=1}^{k} Z_i^2\right) = \sum_{i=1}^{k} E(Z_i^2) = \sum_{i=1}^{k} Var(Z_i) = k.$$

- Hence, from Markov's inequality (1), we get

$$\Pr(Z \geq \varepsilon) = \Pr((\mathbf{X}-\mu)'V^{-1}(\mathbf{X}-\mu) \geq \varepsilon) \leq \frac{E(Z)}{\varepsilon} = \frac{k}{\varepsilon}$$

for all $\varepsilon > 0$.

Another short proof.

- The random vector $\mathbf{Z}$ satisfies $E(\mathbf{Z}) = \mathbf{0}_k$ and

$$Cov(\mathbf{Z}) = Cov(D^{-1/2}T'(\mathbf{X}-\mu)) = D^{-1/2}T'VTD^{-1/2} = D^{-1/2}DD^{-1/2} = I_k.$$

- Therefore $E(Z_i) = 0$, $Var(Z_i) = 1$ and

$$E(Z) = E(\mathbf{Z}'\mathbf{Z}) = E\left(\sum_{i=1}^{k} Z_i^2\right) = \sum_{i=1}^{k} E(Z_i^2) = \sum_{i=1}^{k} Var(Z_i) = k.$$

- Hence, from Markov's inequality (1), we get

$$\Pr(Z \geq \varepsilon) = \Pr((\mathbf{X}-\mu)'V^{-1}(\mathbf{X}-\mu) \geq \varepsilon) \leq \frac{E(Z)}{\varepsilon} = \frac{k}{\varepsilon}$$

for all $\varepsilon > 0$.

Another short proof.

- The random vector $\mathbf{Z}$ satisfies $E(\mathbf{Z}) = \mathbf{0}_k$ and

$$Cov(\mathbf{Z}) = Cov(D^{-1/2}T'(\mathbf{X}-\boldsymbol{\mu})) = D^{-1/2}T'VTD^{-1/2} = D^{-1/2}DD^{-1/2} = I_k.$$

- Therefore $E(Z_i) = 0$, $Var(Z_i) = 1$ and

$$E(Z) = E(\mathbf{Z}'\mathbf{Z}) = E\left(\sum_{i=1}^{k} Z_i^2\right) = \sum_{i=1}^{k} E(Z_i^2) = \sum_{i=1}^{k} Var(Z_i) = k.$$

- Hence, from Markov's inequality (1), we get

$$\Pr(Z \geq \varepsilon) = \Pr((\mathbf{X}-\boldsymbol{\mu})'V^{-1}(\mathbf{X}-\boldsymbol{\mu}) \geq \varepsilon) \leq \frac{E(Z)}{\varepsilon} = \frac{k}{\varepsilon}$$

for all $\varepsilon > 0$.

Bounds for singular covariance matrices.

- $\mathbf{Z} = D^{-1/2}T'(\mathbf{X}-\boldsymbol{\mu})$ is the vector of the standardized principal components of $\mathbf{X}$.
- Then (3) can be written as

$$\Pr(\mathbf{Z}'\mathbf{Z} < \varepsilon) \geq 1 - \frac{k}{\varepsilon} \qquad (7)$$

where $\mathbf{Z}'\mathbf{Z} = \sum_{i=1}^{k} Z_i^2$.
- If V is singular, then $\lambda_1 \geq \cdots \geq \lambda_m > \lambda_{m+1} = \cdots = \lambda_k = 0$.
- Then (7) can be replaced with

$$\Pr\left(\sum_{i=1}^{m} Z_i^2 < \varepsilon\right) \geq 1 - \frac{m}{\varepsilon} \qquad (8)$$

for all $\varepsilon > 0$, where $Z_i = \lambda_i^{-1/2}\mathbf{t}_i'(\mathbf{X}-\mu)$ is the ith standardized principal components of $\mathbf{X}$ and $\mathbf{t}_i$ is the normalized eigenvector associated with the eigenvalue λ_i.

Bounds for singular covariance matrices.

- $\mathbf{Z} = D^{-1/2} T'(\mathbf{X} - \boldsymbol{\mu})$ is the vector of the standardized principal components of $\mathbf{X}$.
- Then (3) can be written as

$$\Pr(\mathbf{Z}'\mathbf{Z} < \varepsilon) \geq 1 - \frac{k}{\varepsilon} \qquad (7)$$

where $\mathbf{Z}'\mathbf{Z} = \sum_{i=1}^{k} Z_i^2$.

- If V is singular, then $\lambda_1 \geq \cdots \geq \lambda_m > \lambda_{m+1} = \cdots = \lambda_k = 0$.
- Then (7) can be replaced with

$$\Pr\left(\sum_{i=1}^{m} Z_i^2 < \varepsilon\right) \geq 1 - \frac{m}{\varepsilon} \qquad (8)$$

for all $\varepsilon > 0$, where $Z_i = \lambda_i^{-1/2}\mathbf{t}_i'(\mathbf{X} - \mu)$ is the ith standardized principal components of $\mathbf{X}$ and $\mathbf{t}_i$ is the normalized eigenvector associated with the eigenvalue λ_i.

Bounds for singular covariance matrices.

- $\mathbf{Z} = D^{-1/2}T'(\mathbf{X} - \boldsymbol{\mu})$ is the vector of the standardized principal components of $\mathbf{X}$.
- Then (3) can be written as

$$\Pr(\mathbf{Z}'\mathbf{Z} < \varepsilon) \geq 1 - \frac{k}{\varepsilon} \quad (7)$$

where $\mathbf{Z}'\mathbf{Z} = \sum_{i=1}^{k} Z_i^2$.

- If V is singular, then $\lambda_1 \geq \cdots \geq \lambda_m > \lambda_{m+1} = \cdots = \lambda_k = 0$.
- Then (7) can be replaced with

$$\Pr\left(\sum_{i=1}^{m} Z_i^2 < \varepsilon\right) \geq 1 - \frac{m}{\varepsilon} \quad (8)$$

for all $\varepsilon > 0$, where $Z_i = \lambda_i^{-1/2}\mathbf{t}_i'(\mathbf{X} - \mu)$ is the ith standardized principal components of $\mathbf{X}$ and $\mathbf{t}_i$ is the normalized eigenvector associated with the eigenvalue λ_i.

Bounds for singular covariance matrices.

- $\mathbf{Z} = D^{-1/2} T'(\mathbf{X} - \boldsymbol{\mu})$ is the vector of the standardized principal components of $\mathbf{X}$.
- Then (3) can be written as

$$\Pr(\mathbf{Z}'\mathbf{Z} < \varepsilon) \geq 1 - \frac{k}{\varepsilon} \tag{7}$$

where $\mathbf{Z}'\mathbf{Z} = \sum_{i=1}^{k} Z_i^2$.
- If V is singular, then $\lambda_1 \geq \cdots \geq \lambda_m > \lambda_{m+1} = \cdots = \lambda_k = 0$.
- Then (7) can be replaced with

$$\Pr\left(\sum_{i=1}^{m} Z_i^2 < \varepsilon\right) \geq 1 - \frac{m}{\varepsilon} \tag{8}$$

for all $\varepsilon > 0$, where $Z_i = \lambda_i^{-1/2} \mathbf{t}_i'(\mathbf{X} - \mu)$ is the ith standardized principal components of $\mathbf{X}$ and $\mathbf{t}_i$ is the normalized eigenvector associated with the eigenvalue λ_i.

An example.

- $(X_1, X_3, X_3) \equiv Multinomial(p_1 = 1/3, p_2 = 1/3, p_3 = 1/3, n)$.
- Then $\mu = E(\mathbf{X}) = (n/3, n/3, n/3)'$ and

$$V = \frac{n}{9} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

- As $X_1 + X_2 + X_3 = n$, we of course have $|V| = 0$,
- The eigenvalues are $\lambda_1 = \lambda_2 = n/3$ and $\lambda_3 = 0$.
- Some two first standardized principal components are

$$Z_1 = \frac{X_1 - X_2}{\sqrt{2n/3}}, \quad Z_2 = \frac{X_1 + X_2 - 2X_3}{\sqrt{2n}}$$

and the multivariate Chebyshev's inequality given in (8) gives

$$\Pr\left(\sqrt{(X_1 - X_2)^2 + (X_1 + X_2 - 2X_3)^2/3} < \delta\right) \geq 1 - \frac{4n}{3\delta^2}.$$

An example.

- $(X_1, X_3, X_3) \equiv Multinomial(p_1 = 1/3, p_2 = 1/3, p_3 = 1/3, n)$.
- Then $\mu = E(\mathbf{X}) = (n/3, n/3, n/3)'$ and

$$V = \frac{n}{9} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

- As $X_1 + X_2 + X_3 = n$, we of course have $|V| = 0$,
- The eigenvalues are $\lambda_1 = \lambda_2 = n/3$ and $\lambda_3 = 0$.
- Some two first standardized principal components are

$$Z_1 = \frac{X_1 - X_2}{\sqrt{2n/3}}, \quad Z_2 = \frac{X_1 + X_2 - 2X_3}{\sqrt{2n}}$$

and the multivariate Chebyshev's inequality given in (8) gives

$$\Pr\left(\sqrt{(X_1 - X_2)^2 + (X_1 + X_2 - 2X_3)^2/3} < \delta\right) \geq 1 - \frac{4n}{3\delta^2}.$$

An example.

- $(X_1, X_3, X_3) \equiv Multinomial(p_1 = 1/3, p_2 = 1/3, p_3 = 1/3, n)$.
- Then $\mu = E(\mathbf{X}) = (n/3, n/3, n/3)'$ and

$$V = \frac{n}{9}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

- As $X_1 + X_2 + X_3 = n$, we of course have $|V| = 0$,
- The eigenvalues are $\lambda_1 = \lambda_2 = n/3$ and $\lambda_3 = 0$.
- Some two first standardized principal components are

$$Z_1 = \frac{X_1 - X_2}{\sqrt{2n/3}}, \quad Z_2 = \frac{X_1 + X_2 - 2X_3}{\sqrt{2n}}$$

and the multivariate Chebyshev's inequality given in (8) gives

$$\Pr\left(\sqrt{(X_1 - X_2)^2 + (X_1 + X_2 - 2X_3)^2/3} < \delta\right) \geq 1 - \frac{4n}{3\delta^2}.$$

An example.

- $(X_1, X_3, X_3) \equiv Multinomial(p_1 = 1/3, p_2 = 1/3, p_3 = 1/3, n)$.
- Then $\mu = E(\mathbf{X}) = (n/3, n/3, n/3)'$ and

$$V = \frac{n}{9} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

- As $X_1 + X_2 + X_3 = n$, we of course have $|V| = 0$,
- The eigenvalues are $\lambda_1 = \lambda_2 = n/3$ and $\lambda_3 = 0$.
- Some two first standardized principal components are

$$Z_1 = \frac{X_1 - X_2}{\sqrt{2n/3}}, \quad Z_2 = \frac{X_1 + X_2 - 2X_3}{\sqrt{2n}}$$

and the multivariate Chebyshev's inequality given in (8) gives

$$\Pr\left(\sqrt{(X_1 - X_2)^2 + (X_1 + X_2 - 2X_3)^2/3} < \delta\right) \geq 1 - \frac{4n}{3\delta^2}.$$

An example.

- $(X_1, X_3, X_3) \equiv Multinomial(p_1 = 1/3, p_2 = 1/3, p_3 = 1/3, n)$.
- Then $\mu = E(\mathbf{X}) = (n/3, n/3, n/3)'$ and

$$V = \frac{n}{9} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

- As $X_1 + X_2 + X_3 = n$, we of course have $|V| = 0$,
- The eigenvalues are $\lambda_1 = \lambda_2 = n/3$ and $\lambda_3 = 0$.
- Some two first standardized principal components are

$$Z_1 = \frac{X_1 - X_2}{\sqrt{2n/3}}, \quad Z_2 = \frac{X_1 + X_2 - 2X_3}{\sqrt{2n}}$$

and the multivariate Chebyshev's inequality given in (8) gives

$$\Pr\left(\sqrt{(X_1 - X_2)^2 + (X_1 + X_2 - 2X_3)^2/3} < \delta\right) \geq 1 - \frac{4n}{3\delta^2}.$$

The bounds are sharp.

Theorem (Navarro SPL 2014)

Let $\mathbf{X} = (X_1, \ldots, X_k)'$ *be a random vector with finite mean vector* $\mu = E(\mathbf{X})$ *and positive definite covariance matrix* $V = Cov(\mathbf{X})$ *and let* $\varepsilon \geq k$*. Then there exists a sequence* $\mathbf{X}^{(n)} = (X_1^{(n)}, \ldots, X_k^{(n)})'$ *of random vectors with mean vector* μ *and covariance matrix* V *such that*

$$\lim_{n\to\infty} \Pr((\mathbf{X}^{(n)} - \mu)' V^{-1} (\mathbf{X}^{(n)} - \mu) \geq \varepsilon) = \frac{k}{\varepsilon}. \qquad (9)$$

The bounds are sharp (proof).

- For $\varepsilon \geq k$, let us consider

$$D_n = \begin{cases} \sqrt{Z_n + \varepsilon} & \text{with probability } (p - 1/n)/2 \\ -\sqrt{Z_n + \varepsilon} & \text{with probability } (p - 1/n)/2 \\ 0 & \text{with probability } 1 - p + 1/n \end{cases}$$

  for $n > \varepsilon/k$, where $p = k/\varepsilon \leq 1$ and
  $Z_n \equiv Exp(\mu_n = \frac{\varepsilon/n}{p-1/n} > 0)$.

- Note that $\Pr(D_n^2 \geq \varepsilon) = p - 1/n$.
- $E(D_n) = \frac{(p-1/n)}{2} E\left(\sqrt{Z_n + \varepsilon}\right) - \frac{(p-1/n)}{2} E\left(\sqrt{Z_n + \varepsilon}\right) = 0.$
- $E(D_n^2) = (p - 1/n)E(Z_n + \varepsilon) = (p - 1/n)\left(\frac{\varepsilon/n}{p-1/n} + \varepsilon\right) = k.$

The bounds are sharp (proof).

- For $\varepsilon \geq k$, let us consider

$$D_n = \begin{cases} \sqrt{Z_n + \varepsilon} & \text{with probability } (p - 1/n)/2 \\ -\sqrt{Z_n + \varepsilon} & \text{with probability } (p - 1/n)/2 \\ 0 & \text{with probability } 1 - p + 1/n \end{cases}$$

  for $n > \varepsilon/k$, where $p = k/\varepsilon \leq 1$ and $Z_n \equiv Exp(\mu_n = \frac{\varepsilon/n}{p-1/n} > 0)$.
- Note that $\Pr(D_n^2 \geq \varepsilon) = p - 1/n$.
- $E(D_n) = \frac{(p-1/n)}{2} E\left(\sqrt{Z_n + \varepsilon}\right) - \frac{(p-1/n)}{2} E\left(\sqrt{Z_n + \varepsilon}\right) = 0.$
- $E(D_n^2) = (p - 1/n)E(Z_n + \varepsilon) = (p - 1/n)\left(\frac{\varepsilon/n}{p-1/n} + \varepsilon\right) = k.$

The bounds are sharp (proof).

- For $\varepsilon \geq k$, let us consider

$$D_n = \begin{cases} \sqrt{Z_n + \varepsilon} & \text{with probability } (p - 1/n)/2 \\ -\sqrt{Z_n + \varepsilon} & \text{with probability } (p - 1/n)/2 \\ 0 & \text{with probability } 1 - p + 1/n \end{cases}$$

  for $n > \varepsilon/k$, where $p = k/\varepsilon \leq 1$ and $Z_n \equiv Exp(\mu_n = \frac{\varepsilon/n}{p-1/n} > 0)$.

- Note that $\Pr(D_n^2 \geq \varepsilon) = p - 1/n$.
- $E(D_n) = \frac{(p-1/n)}{2} E\left(\sqrt{Z_n + \varepsilon}\right) - \frac{(p-1/n)}{2} E\left(\sqrt{Z_n + \varepsilon}\right) = 0.$
- $E(D_n^2) = (p - 1/n)E(Z_n + \varepsilon) = (p - 1/n)\left(\frac{\varepsilon/n}{p-1/n} + \varepsilon\right) = k.$

The bounds are sharp (proof).

- For $\varepsilon \geq k$, let us consider

$$D_n = \begin{cases} \sqrt{Z_n + \varepsilon} & \text{with probability } (p - 1/n)/2 \\ -\sqrt{Z_n + \varepsilon} & \text{with probability } (p - 1/n)/2 \\ 0 & \text{with probability } 1 - p + 1/n \end{cases}$$

  for $n > \varepsilon/k$, where $p = k/\varepsilon \leq 1$ and $Z_n \equiv Exp(\mu_n = \frac{\varepsilon/n}{p-1/n} > 0)$.

- Note that $\Pr(D_n^2 \geq \varepsilon) = p - 1/n$.
- $E(D_n) = \frac{(p-1/n)}{2} E\left(\sqrt{Z_n + \varepsilon}\right) - \frac{(p-1/n)}{2} E\left(\sqrt{Z_n + \varepsilon}\right) = 0.$
- $E(D_n^2) = (p - 1/n)E(Z_n + \varepsilon) = (p - 1/n)\left(\frac{\varepsilon/n}{p-1/n} + \varepsilon\right) = k.$

J. Navarro, E-mail: jorgenav@um.es

The bounds are sharp (proof).

- Let U_n be a r.v., independent of Z_n, with a uniform distribution over $\{1,\dots,k\}$.
- Let $\mathbf{Y}^{(n)} = (Y_1^{(n)},\dots,Y_k^{(n)})'$ defined by $Y_i^{(n)} = D_n$ and $Y_j^{(n)} = 0$ for $j = 1,\dots,i-1,i+1,\dots,k$ when $U_n = i$.
- Hence $E(Y_i^{(n)}) = \frac{1}{k}E(D_n) = 0$ and

$$Var(Y_i^{(n)}) = E((Y_i^{(n)})^2) = \frac{1}{k}E(D_n^2) = 1.$$

- Moreover, $Y_i^{(n)}Y_j^{(n)} = 0$ and $E(Y_i^{(n)}Y_j^{(n)}) = 0$ for all $i \neq j$.
- Then $E(\mathbf{Y}^{(n)}) = \mathbf{0}_k$ and $Cov(\mathbf{Y}^{(n)}) = I_k$.
- Then $\mathbf{X}^{(n)} = \mu + V^{1/2}\mathbf{Y}^{(n)}$ has mean $E(\mathbf{X}^{(n)}) = \mu$ and

$$Cov(\mathbf{X}^{(n)}) = Cov(V^{1/2}\mathbf{Y}^{(n)}) = V^{1/2}V^{1/2} = V.$$

The bounds are sharp (proof).

- Let U_n be a r.v., independent of Z_n, with a uniform distribution over $\{1,\ldots,k\}$.
- Let $\mathbf{Y}^{(n)} = (Y_1^{(n)},\ldots,Y_k^{(n)})'$ defined by $Y_i^{(n)} = D_n$ and $Y_j^{(n)} = 0$ for $j = 1,\ldots,i-1,i+1,\ldots,k$ when $U_n = i$.
- Hence $E(Y_i^{(n)}) = \frac{1}{k}E(D_n) = 0$ and

$$Var(Y_i^{(n)}) = E((Y_i^{(n)})^2) = \frac{1}{k}E(D_n^2) = 1.$$

- Moreover, $Y_i^{(n)}Y_j^{(n)} = 0$ and $E(Y_i^{(n)}Y_j^{(n)}) = 0$ for all $i \neq j$.
- Then $E(\mathbf{Y}^{(n)}) = \mathbf{0}_k$ and $Cov(\mathbf{Y}^{(n)}) = I_k$.
- Then $\mathbf{X}^{(n)} = \mu + V^{1/2}\mathbf{Y}^{(n)}$ has mean $E(\mathbf{X}^{(n)}) = \mu$ and

$$Cov(\mathbf{X}^{(n)}) = Cov(V^{1/2}\mathbf{Y}^{(n)}) = V^{1/2}V^{1/2} = V.$$

The bounds are sharp (proof).

- Let U_n be a r.v., independent of Z_n, with a uniform distribution over $\{1,\ldots,k\}$.
- Let $\mathbf{Y}^{(n)} = (Y_1^{(n)},\ldots,Y_k^{(n)})'$ defined by $Y_i^{(n)} = D_n$ and $Y_j^{(n)} = 0$ for $j = 1,\ldots,i-1,i+1,\ldots,k$ when $U_n = i$.
- Hence $E(Y_i^{(n)}) = \frac{1}{k}E(D_n) = 0$ and

$$Var(Y_i^{(n)}) = E((Y_i^{(n)})^2) = \frac{1}{k}E(D_n^2) = 1.$$

- Moreover, $Y_i^{(n)}Y_j^{(n)} = 0$ and $E(Y_i^{(n)}Y_j^{(n)}) = 0$ for all $i \neq j$.
- Then $E(\mathbf{Y}^{(n)}) = \mathbf{0}_k$ and $Cov(\mathbf{Y}^{(n)}) = I_k$.
- Then $\mathbf{X}^{(n)} = \mu + V^{1/2}\mathbf{Y}^{(n)}$ has mean $E(\mathbf{X}^{(n)}) = \mu$ and

$$Cov(\mathbf{X}^{(n)}) = Cov(V^{1/2}\mathbf{Y}^{(n)}) = V^{1/2}V^{1/2} = V.$$

The bounds are sharp (proof).

- Let U_n be a r.v., independent of Z_n, with a uniform distribution over $\{1, \ldots, k\}$.
- Let $\mathbf{Y}^{(n)} = (Y_1^{(n)}, \ldots, Y_k^{(n)})'$ defined by $Y_i^{(n)} = D_n$ and $Y_j^{(n)} = 0$ for $j = 1, \ldots, i-1, i+1, \ldots, k$ when $U_n = i$.
- Hence $E(Y_i^{(n)}) = \frac{1}{k} E(D_n) = 0$ and

$$Var(Y_i^{(n)}) = E((Y_i^{(n)})^2) = \frac{1}{k} E(D_n^2) = 1.$$

- Moreover, $Y_i^{(n)} Y_j^{(n)} = 0$ and $E(Y_i^{(n)} Y_j^{(n)}) = 0$ for all $i \neq j$.
- Then $E(\mathbf{Y}^{(n)}) = \mathbf{0}_k$ and $Cov(\mathbf{Y}^{(n)}) = I_k$.
- Then $\mathbf{X}^{(n)} = \mu + V^{1/2} \mathbf{Y}^{(n)}$ has mean $E(\mathbf{X}^{(n)}) = \mu$ and

$$Cov(\mathbf{X}^{(n)}) = Cov(V^{1/2} \mathbf{Y}^{(n)}) = V^{1/2} V^{1/2} = V.$$

The bounds are sharp (proof).

- Let U_n be a r.v., independent of Z_n, with a uniform distribution over $\{1,\dots,k\}$.
- Let $\mathbf{Y}^{(n)} = (Y_1^{(n)},\dots,Y_k^{(n)})'$ defined by $Y_i^{(n)} = D_n$ and $Y_j^{(n)} = 0$ for $j = 1,\dots,i-1,i+1,\dots,k$ when $U_n = i$.
- Hence $E(Y_i^{(n)}) = \frac{1}{k}E(D_n) = 0$ and

$$Var(Y_i^{(n)}) = E((Y_i^{(n)})^2) = \frac{1}{k}E(D_n^2) = 1.$$

- Moreover, $Y_i^{(n)}Y_j^{(n)} = 0$ and $E(Y_i^{(n)}Y_j^{(n)}) = 0$ for all $i \neq j$.
- Then $E(\mathbf{Y}^{(n)}) = \mathbf{0}_k$ and $Cov(\mathbf{Y}^{(n)}) = I_k$.
- Then $\mathbf{X}^{(n)} = \mu + V^{1/2}\mathbf{Y}^{(n)}$ has mean $E(\mathbf{X}^{(n)}) = \mu$ and

$$Cov(\mathbf{X}^{(n)}) = Cov(V^{1/2}\mathbf{Y}^{(n)}) = V^{1/2}V^{1/2} = V.$$

The bounds are sharp (proof).

- Let U_n be a r.v., independent of Z_n, with a uniform distribution over $\{1, \dots, k\}$.
- Let $\mathbf{Y}^{(n)} = (Y_1^{(n)}, \dots, Y_k^{(n)})'$ defined by $Y_i^{(n)} = D_n$ and $Y_j^{(n)} = 0$ for $j = 1, \dots, i-1, i+1, \dots, k$ when $U_n = i$.
- Hence $E(Y_i^{(n)}) = \frac{1}{k} E(D_n) = 0$ and

$$Var(Y_i^{(n)}) = E((Y_i^{(n)})^2) = \frac{1}{k} E(D_n^2) = 1.$$

- Moreover, $Y_i^{(n)} Y_j^{(n)} = 0$ and $E(Y_i^{(n)} Y_j^{(n)}) = 0$ for all $i \neq j$.
- Then $E(\mathbf{Y}^{(n)}) = \mathbf{0}_k$ and $Cov(\mathbf{Y}^{(n)}) = I_k$.
- Then $\mathbf{X}^{(n)} = \mu + V^{1/2} \mathbf{Y}^{(n)}$ has mean $E(\mathbf{X}^{(n)}) = \mu$ and

$$Cov(\mathbf{X}^{(n)}) = Cov(V^{1/2} \mathbf{Y}^{(n)}) = V^{1/2} V^{1/2} = V.$$

The bounds are sharp (proof).

- Moreover,

$$\begin{aligned}
\Pr((\mathbf{X}^{(n)}-\mu)'V^{-1}(\mathbf{X}^{(n)}-\mu)\geq\varepsilon) \\
&= \Pr((V^{1/2}\mathbf{Y}^{(n)})'V^{-1}(V^{1/2}\mathbf{Y}^{(n)})\geq\varepsilon) \\
&= \Pr((\mathbf{Y}^{(n)})'V^{1/2}V^{-1}V^{1/2}\mathbf{Y}^{(n)}\geq\varepsilon) \\
&= \Pr((\mathbf{Y}^{(n)})'\mathbf{Y}^{(n)}\geq\varepsilon) \\
&= \Pr\left(\sum_{i=1}^{k}(Y_i^{(n)})^2\geq\varepsilon\right) \\
&= \Pr(D_n^2\geq\varepsilon) \\
&= p-\frac{1}{n}\to p=\frac{k}{\varepsilon},\ \text{as } n\to\infty
\end{aligned}$$

Applications. Case $k = 2$.

Theorem

$(X, Y)'$ *with* $E(X) = \mu_X$, $E(Y) = \mu_Y$, $Var(X) = \sigma_X^2 > 0$, $Var(Y) = \sigma_Y^2 > 0$ *and* $\rho = Cor(X, Y) \in (-1, 1)$. *Then*

$$\Pr((X^* - Y^*)^2 + 2(1-\rho)X^*Y^* < \delta) \geq 1 - 2\frac{1-\rho^2}{\delta} \qquad (10)$$

for all $\delta > 0$, *where* $X^* = (X - \mu_X)/\sigma_X$ *and* $Y^* = (X - \mu_Y)/\sigma_Y$.

$Z_1 = (X^* + Y^*)/\sqrt{2(1+\rho)}$, $Z_2 = (X^* - Y^*)/\sqrt{2(1-\rho)}$ and

$$\Pr\left(\frac{(X^*+Y^*)^2}{2(1+\rho)} + \frac{(X^*-Y^*)^2}{2(1-\rho)} < \varepsilon\right) \geq 1 - \frac{2}{\varepsilon}. \qquad (11)$$

An example

- (X, Y) with $E(X) = E(Y) = 1$, $Var(X) = Var(Y) = 1$ and $\rho = Cor(X, Y) = 0.9$. Then

$$\Pr(5(X-Y)^2 + (X-1)(Y-1) < 5\delta) \geq 1 - 2\frac{0.19}{\delta},$$

that is,

$$\Pr(5X^2 - 9XY + 5Y^2 - X - Y + 1 < \varepsilon) \geq 1 - \frac{1.9}{\varepsilon}$$

for all $\varepsilon > 1.9$.

- The distribution-free confidence regions for $\varepsilon = 3, 4, 5, 10$ containing respectively at least the 36.6666%, 52.5%, 62% and the 81% of the values of (X, Y) can be seen in the following figure.

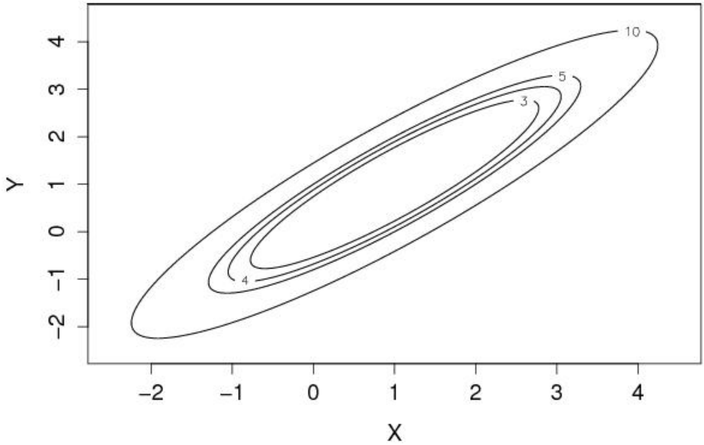

Figure: Confidence regions for $\varepsilon = 3, 4, 5, 10$ containing at least the 36.66%, 52.5%, 62% and the 81% of the values of (X, Y).

Order statistics

- Let $X_{1:k},\dots,X_{k:k}$ be the OS from $(X_1,\dots,X_k)$.
- For $k=2$ we have

$$\rho_{1,2:2}=Cor(X_{1:2},X_{2:2})=\rho\frac{\sigma_1\sigma_2}{\sigma_{1:2}\sigma_{1:2}}+\frac{(\mu_1-\mu_{1:2})(\mu_2-\mu_{1:2})}{\sigma_{1:2}\sigma_{1:2}},$$

where $\mu_i=E(X_i)$, $\mu_{i:2}=E(X_{i:2})$, $\sigma_i^2=Var(X_i)$, $\sigma_{i:2}^2=Var(X_{i:2})$, for $i=1,2$, and $\rho=Cor(X_1,X_2)$.
- Then

$$\Pr((X^*_{2:2}-X^*_{1:2})^2+2(1-\rho_{1,2:2})X^*_{2:2}X^*_{1:2}<\delta)\geq 1-2\frac{1-\rho^2_{1,2:2}}{\delta}, \tag{12}$$

where $X^*_{i:2}=(X_{i:2}-\mu_{i:2})/\sigma_{i:2}$, $i=1,2$.

Order statistics

- Let $X_{1:k}, \dots, X_{k:k}$ be the OS from $(X_1, \dots, X_k)$.
- For $k = 2$ we have

$$\rho_{1,2:2} = Cor(X_{1:2}, X_{2:2}) = \rho \frac{\sigma_1 \sigma_2}{\sigma_{1:2}\sigma_{1:2}} + \frac{(\mu_1 - \mu_{1:2})(\mu_2 - \mu_{1:2})}{\sigma_{1:2}\sigma_{1:2}},$$

where $\mu_i = E(X_i)$, $\mu_{i:2} = E(X_{i:2})$, $\sigma_i^2 = Var(X_i)$, $\sigma_{i:2}^2 = Var(X_{i:2})$, for $i = 1, 2$, and $\rho = Cor(X_1, X_2)$.

- Then

$$\Pr((X^*_{2:2} - X^*_{1:2})^2 + 2(1 - \rho_{1,2:2}) X^*_{2:2} X^*_{1:2} < \delta) \geq 1 - 2 \frac{1 - \rho^2_{1,2:2}}{\delta}, \quad (12)$$

where $X^*_{i:2} = (X_{i:2} - \mu_{i:2})/\sigma_{i:2}$, $i = 1, 2$.

Order statistics

- Let $X_{1:k}, \dots, X_{k:k}$ be the OS from $(X_1, \dots, X_k)$.
- For $k = 2$ we have

$$\rho_{1,2:2} = Cor(X_{1:2}, X_{2:2}) = \rho \frac{\sigma_1 \sigma_2}{\sigma_{1:2}\sigma_{1:2}} + \frac{(\mu_1 - \mu_{1:2})(\mu_2 - \mu_{1:2})}{\sigma_{1:2}\sigma_{1:2}},$$

where $\mu_i = E(X_i)$, $\mu_{i:2} = E(X_{i:2})$, $\sigma_i^2 = Var(X_i)$, $\sigma_{i:2}^2 = Var(X_{i:2})$, for $i = 1, 2$, and $\rho = Cor(X_1, X_2)$.

- Then

$$\Pr((X^*_{2:2} - X^*_{1:2})^2 + 2(1 - \rho_{1,2:2}) X^*_{2:2} X^*_{1:2} < \delta) \geq 1 - 2\frac{1 - \rho^2_{1,2:2}}{\delta}, \tag{12}$$

where $X^*_{i:2} = (X_{i:2} - \mu_{i:2})/\sigma_{i:2}$, $i = 1, 2$.

Order statistics. Example 1.

- (X_1, X_2) has a Pareto distribution with

$$\overline{F}(x,y) = \Pr(X_1 > x, X_2 > y) = (1 + \lambda x + \lambda y)^{-\theta}$$

for $x, y \geq 0$, where $\lambda > 0$ and $\theta > 2$.

- Then $\mu = 1/(\lambda\theta - \lambda)$, $\sigma^2 = \mu^2/(1-2\rho)$, $\rho = 1/\theta$, $\mu_{1:2} = \mu/2$, $\mu_{2:2} = 3\mu/2$

$$\sigma_{1:2}^2 = \frac{\mu^2}{4(1-2\rho)}, \quad \sigma_{2:2}^2 = \frac{\mu^2(6+3\rho)}{4(1-2\rho)}, \quad \rho_{1,2:2} = \frac{1+2\rho}{\sqrt{6+3\rho}}.$$

- If $\lambda = 0.5$ and $\theta = 3$, then $\mu = 1$, $\rho = 1/3$, $\mu_{1:2} = 1/2$, $\mu_{2:2} = 3/2$, $\sigma_{1:2} = 0.866$, $\sigma_{2:2} = 2.291$ $\rho_{1,2:2} = 0.6299$ and

$$\Pr\left(\left[\frac{X_{2:2}-\frac{3}{2}}{2.291} - \frac{X_{1:2}-\frac{1}{2}}{0.866}\right]^2 + 0.74\frac{X_{2:2}-\frac{3}{2}}{2.291}\frac{X_{1:2}-\frac{1}{2}}{0.866} < \delta\right) \geq 1 - \frac{1.206}{\delta}.$$

Order statistics. Example 1.

- (X_1, X_2) has a Pareto distribution with

$$\overline{F}(x,y) = \Pr(X_1 > x, X_2 > y) = (1 + \lambda x + \lambda y)^{-\theta}$$

  for $x, y \geq 0$, where $\lambda > 0$ and $\theta > 2$.
- Then $\mu = 1/(\lambda\theta - \lambda)$, $\sigma^2 = \mu^2/(1-2\rho)$, $\rho = 1/\theta$, $\mu_{1:2} = \mu/2$, $\mu_{2:2} = 3\mu/2$

$$\sigma^2_{1:2} = \frac{\mu^2}{4(1-2\rho)}, \quad \sigma^2_{2:2} = \frac{\mu^2(6+3\rho)}{4(1-2\rho)}, \quad \rho_{1,2:2} = \frac{1+2\rho}{\sqrt{6+3\rho}}.$$

- If $\lambda = 0.5$ and $\theta = 3$, then $\mu = 1$, $\rho = 1/3$, $\mu_{1:2} = 1/2$, $\mu_{2:2} = 3/2$, $\sigma_{1:2} = 0.866$, $\sigma_{2:2} = 2.291$ $\rho_{1,2:2} = 0.6299$ and

$$\Pr\left(\left[\frac{X_{2:2}-\frac{3}{2}}{2.291} - \frac{X_{1:2}-\frac{1}{2}}{0.866}\right]^2 + 0.74\frac{X_{2:2}-\frac{3}{2}}{2.291}\frac{X_{1:2}-\frac{1}{2}}{0.866} < \delta\right) \geq 1 - \frac{1.206}{\delta}.$$

Order statistics. Example 1.

- (X_1, X_2) has a Pareto distribution with

$$\overline{F}(x,y) = \Pr(X_1 > x, X_2 > y) = (1 + \lambda x + \lambda y)^{-\theta}$$

for $x, y \geq 0$, where $\lambda > 0$ and $\theta > 2$.

- Then $\mu = 1/(\lambda\theta - \lambda)$, $\sigma^2 = \mu^2/(1-2\rho)$, $\rho = 1/\theta$, $\mu_{1:2} = \mu/2$, $\mu_{2:2} = 3\mu/2$

$$\sigma^2_{1:2} = \frac{\mu^2}{4(1-2\rho)}, \quad \sigma^2_{2:2} = \frac{\mu^2(6+3\rho)}{4(1-2\rho)}, \quad \rho_{1,2:2} = \frac{1+2\rho}{\sqrt{6+3\rho}}.$$

- If $\lambda = 0.5$ and $\theta = 3$, then $\mu = 1$, $\rho = 1/3$, $\mu_{1:2} = 1/2$, $\mu_{2:2} = 3/2$, $\sigma_{1:2} = 0.866$, $\sigma_{2:2} = 2.291$ $\rho_{1,2:2} = 0.6299$ and

$$\Pr\left(\left[\frac{X_{2:2} - \frac{3}{2}}{2.291} - \frac{X_{1:2} - \frac{1}{2}}{0.866}\right]^2 + 0.74\frac{X_{2:2} - \frac{3}{2}}{2.291}\frac{X_{1:2} - \frac{1}{2}}{0.866} < \delta\right) \geq 1 - \frac{1.206}{\delta}.$$

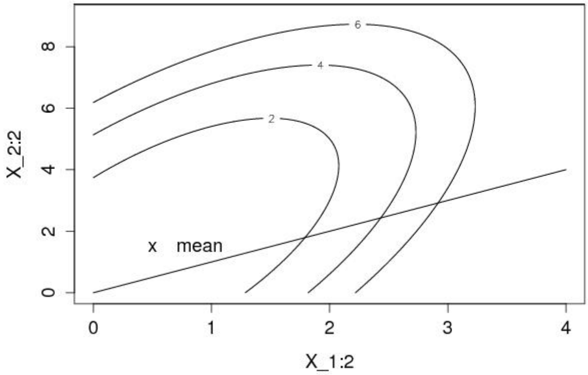

Figure: Confidence regions for $\delta = 2, 4, 6$ containing at least the 39.68%, the 69.84% and the 79.89% of the values of $(X_{1:2}, X_{2:2})$.

Order statistics. Example 2.

- $X_1, \dots, X_k$ iid $Exp(\mu = 1)$, then

$$\mu_{i:k} = \sum_{j=k-i+1}^{k} \frac{1}{j}, \quad \sigma^2_{i:k} = \sum_{j=k-i+1}^{k} \frac{1}{j^2}$$

and

$$\rho_{i,j:k} = Cor(X_{i:k}, X_{j:k}) = \frac{\sigma_{i:k}}{\sigma_{j:k}}, \quad 1 \le i < j \le k$$

- If $k = 3$, $i = 2$ and $j = 3$, then $\mu_{2:3} = 5/6$, $\mu_{3:3} = 11/6$, $\sigma_{2:3} = 0.6009$, $\sigma_{3:3} = 1.1667$, and $\rho_{2,3:3} = 0.5151$.
- Hence

$$\Pr\left(\left[\frac{X_{3:3} - \frac{11}{6}}{1.1667} - \frac{X_{2:3} - \frac{5}{6}}{0.6009}\right]^2 + 0.969 \frac{X_{3:3} - \frac{11}{6}}{1.1667} \frac{X_{2:3} - \frac{5}{6}}{0.6009} < \delta\right) \ge 1 - \frac{1.469}{\delta}.$$

Order statistics. Example 2.

- $X_1,\dots,X_k$ iid $Exp(\mu=1)$, then

$$\mu_{i:k}=\sum_{j=k-i+1}^{k}\frac{1}{j},\quad \sigma^2_{i:k}=\sum_{j=k-i+1}^{k}\frac{1}{j^2}$$

and

$$\rho_{i,j:k}=Cor(X_{i:k},X_{j:k})=\frac{\sigma_{i:k}}{\sigma_{j:k}},\quad 1\le i<j\le k$$

- If $k=3$, $i=2$ and $j=3$, then $\mu_{2:3}=5/6$, $\mu_{3:3}=11/6$, $\sigma_{2:3}=0.6009$, $\sigma_{3:3}=1.1667$, and $\rho_{2,3:3}=0.5151$.
- Hence

$$\Pr\left(\left[\frac{X_{3:3}-\frac{11}{6}}{1.1667}-\frac{X_{2:3}-\frac{5}{6}}{0.6009}\right]^2+0.969\frac{X_{3:3}-\frac{11}{6}}{1.1667}\frac{X_{2:3}-\frac{5}{6}}{0.6009}<\delta\right)\ge 1-\frac{1.469}{\delta}.$$

Order statistics. Example 2.

- $X_1, \ldots, X_k$ iid $Exp(\mu = 1)$, then

$$\mu_{i:k} = \sum_{j=k-i+1}^{k} \frac{1}{j}, \quad \sigma^2_{i:k} = \sum_{j=k-i+1}^{k} \frac{1}{j^2}$$

and

$$\rho_{i,j:k} = Cor(X_{i:k}, X_{j:k}) = \frac{\sigma_{i:k}}{\sigma_{j:k}}, \quad 1 \leq i < j \leq k$$

- If $k = 3$, $i = 2$ and $j = 3$, then $\mu_{2:3} = 5/6$, $\mu_{3:3} = 11/6$, $\sigma_{2:3} = 0.6009$, $\sigma_{3:3} = 1.1667$, and $\rho_{2,3:3} = 0.5151$.
- Hence

$$\Pr\left(\left[\frac{X_{3:3} - \frac{11}{6}}{1.1667} - \frac{X_{2:3} - \frac{5}{6}}{0.6009}\right]^2 + 0.969\frac{X_{3:3} - \frac{11}{6}}{1.1667}\frac{X_{2:3} - \frac{5}{6}}{0.6009} < \delta\right) \geq 1 - \frac{1.469}{\delta}.$$

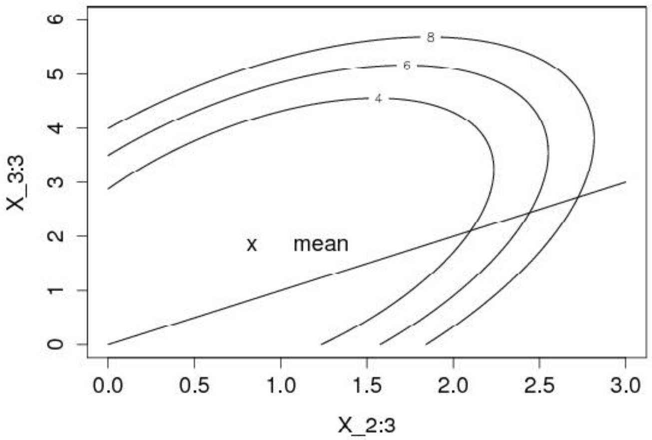

Figure: Confidence regions for $\delta = 2, 3, 4$ containing at least 63.26%, the 75.51% and the 81.63% of the values of $(X_{2:3}, X_{3:3})$.

Order statistics. Example 2.

- For $(X_{1:3}, X_{2:3}, X_{3:3})'$ we obtain the confidence region

$$R_\varepsilon = \{(x, y, z) : 1.444x^2 - 1.602xy + 1.805y^2 - 1.402yz + 1.361z^2 < \varepsilon\}$$

containing $(X^*_{1:3}, X^*_{2:3}, X^*_{3:3})'$ with a probability greater than $1 - 3/\varepsilon$, where $X^*_{i:k} = (X_{i:k} - \mu_{i:k})/\sigma_{i:k}$ for $i = 1, 2, 3$.

- If we use the two principal components

$$\Pr\left(\frac{Y_1^2}{1.9129431} + \frac{Y_2^2}{0.77153779} < \varepsilon\right) \geq 1 - \frac{2}{\varepsilon} \qquad (13)$$

for all $\varepsilon > 0$, where

$$Y_1 = 0.5548133X^*_{1:3} + 0.6382230X^*_{2:3} + 0.5337169X^*_{3:3}$$

and

$$Y_2 = 0.66914423X^*_{1:3} + 0.03890251X^*_{2:3} - 0.7421136X^*_{3:3}.$$

Order statistics. Example 2.

- For $(X_{1:3}, X_{2:3}, X_{3:3})'$ we obtain the confidence region

$$R_\varepsilon = \{(x, y, z) : 1.444x^2 - 1.602xy + 1.805y^2 - 1.402yz + 1.361z^2 < \varepsilon\}$$

containing $(X^*_{1:3}, X^*_{2:3}, X^*_{3:3})'$ with a probability greater than $1 - 3/\varepsilon$, where $X^*_{i:k} = (X_{i:k} - \mu_{i:k})/\sigma_{i:k}$ for $i = 1, 2, 3$.

- If we use the two principal components

$$\Pr\left(\frac{Y_1^2}{1.9129431} + \frac{Y_2^2}{0.77153779} < \varepsilon\right) \geq 1 - \frac{2}{\varepsilon} \qquad (13)$$

for all $\varepsilon > 0$, where

$$Y_1 = 0.5548133X^*_{1:3} + 0.6382230X^*_{2:3} + 0.5337169X^*_{3:3}$$

and

$$Y_2 = 0.66914423X^*_{1:3} + 0.03890251X^*_{2:3} - 0.7421136X^*_{3:3}.$$

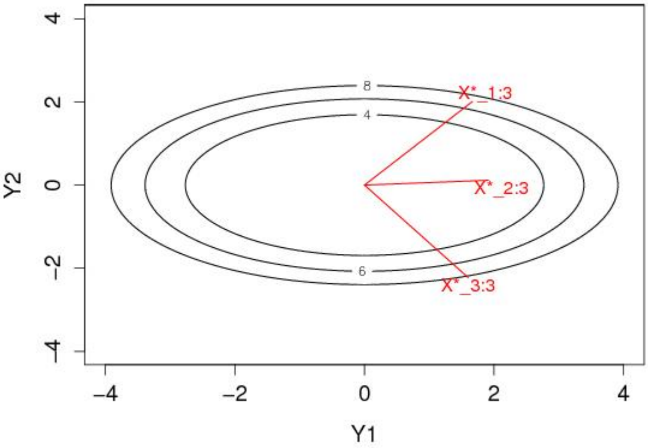

Figure: Confidence regions for $\varepsilon = 4, 6, 8$ containing at least the 50%, the 66.6667% and the 75% of the scores of $(X_{1:3}, X_{2:3}, X_{3:3})$.

Data sets.

- If we have a data set $O_i = (X_i, Y_i)'$, $i = 1, \dots, n$, the mean is

$$\overline{O} = \frac{1}{n} \sum_{i=1}^{n} O_i = (\overline{X}, \overline{Y})$$

and its covariance matrix is

$$\widehat{V} = \frac{1}{n} \sum_{m=1}^{n} (O_m - \overline{O})(O_m - \overline{O})' = (\widehat{V}_{i,j}),$$

- The correlation is $r = \widehat{V}_{1,2} / \sqrt{\widehat{V}_{1,1} \widehat{V}_{2,2}}$ and

$$\Pr((X_I^* - Y_I^*)^2 + 2(1-r)X_I^* Y_I^* < \delta) \geq 1 - 2\frac{1-r^2}{\delta}, \quad (14)$$

where $X_I^* = (X_I - \overline{X})/\sqrt{\widehat{V}_{1,1}}$, $Y_I^* = (Y_I - \overline{Y})/\sqrt{\widehat{V}_{2,2}}$ and $I = i$ with probability $1/n$.

Data sets.

- If we have a data set $O_i = (X_i, Y_i)'$, $i = 1, \dots, n$, the mean is

$$\overline{O} = \frac{1}{n} \sum_{i=1}^{n} O_i = (\overline{X}, \overline{Y})$$

and its covariance matrix is

$$\widehat{V} = \frac{1}{n} \sum_{m=1}^{n} (O_m - \overline{O})(O_m - \overline{O})' = (\widehat{V}_{i,j}),$$

- The correlation is $r = \widehat{V}_{1,2} / \sqrt{\widehat{V}_{1,1} \widehat{V}_{2,2}}$ and

$$\Pr((X_I^* - Y_I^*)^2 + 2(1-r) X_I^* Y_I^* < \delta) \geq 1 - 2 \frac{1 - r^2}{\delta}, \quad (14)$$

where $X_I^* = (X_I - \overline{X}) / \sqrt{\widehat{V}_{1,1}}$, $Y_I^* = (Y_I - \overline{Y}) / \sqrt{\widehat{V}_{2,2}}$ and $I = i$ with probability $1/n$.

Data sets.

- Then, by taking $\delta = 4(1 - r^2)$

$$R_1 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^*y^* < 4(1 - r^2)\},$$

contains (for sure) at least the 50% of the data.

- By taking $\delta = 8(1 - r^2)$

$$R_2 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^*y^* < 8(1 - r^2)\},$$

contains (for sure) at least the 75% of the data and the complementary region

$$\overline{R}_2 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^*y^* \geq 8(1 - r^2)\},$$

contains (for sure) at most the 25% of the data.

- These regions are similar to (univariate) box plots.

Data sets.

- Then, by taking $\delta = 4(1 - r^2)$

$$R_1 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^* y^* < 4(1 - r^2)\},$$

contains (for sure) at least the 50% of the data.

- By taking $\delta = 8(1 - r^2)$

$$R_2 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^* y^* < 8(1 - r^2)\},$$

contains (for sure) at least the 75% of the data and the complementary region

$$\overline{R}_2 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^* y^* \geq 8(1 - r^2)\},$$

contains (for sure) at most the 25% of the data.

- These regions are similar to (univariate) box plots.

Data sets.

- Then, by taking $\delta = 4(1 - r^2)$

$$R_1 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^*y^* < 4(1 - r^2)\},$$

contains (for sure) at least the 50% of the data.

- By taking $\delta = 8(1 - r^2)$

$$R_2 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^*y^* < 8(1 - r^2)\},$$

contains (for sure) at least the 75% of the data and the complementary region

$$\overline{R}_2 = \{(x, y) \in \mathbb{R}^2 : (x^* - y^*)^2 + 2(1 - r)x^*y^* \geq 8(1 - r^2)\},$$

contains (for sure) at most the 25% of the data.

- These regions are similar to (univariate) box plots.

Data sets. An example.

- Consider in the data set "*iris*" from R (Fisher, 1936), the variables $X = Petal.Length$ and $Y = Petal.Width$.
- We obtain $r = 0.9628654$ and R_1 and R_2 determined by

$$\left(\frac{x-3.758}{1.759}-\frac{y-1.199}{0.759}\right)^2+2(1-r)\frac{x-3.758}{1.759}\frac{y-1.199}{0.759}<0.292$$

and

$$\left(\frac{x-3.758}{1.759}-\frac{y-1.199}{0.759}\right)^2+2(1-r)\frac{x-3.758}{1.759}\frac{y-1.199}{0.759}<0.583,$$

respectively.
- These regions contain more than the 50% and the 75% of the data (i.e. more than 75 and 113 data in this case).

Data sets. An example.

- Consider in the data set "*iris*" from R (Fisher, 1936), the variables $X = Petal.Length$ and $Y = Petal.Width$.
- We obtain $r = 0.9628654$ and R_1 and R_2 determined by

$$\left(\frac{x-3.758}{1.759}-\frac{y-1.199}{0.759}\right)^2+2(1-r)\frac{x-3.758}{1.759}\frac{y-1.199}{0.759}<0.292$$

and

$$\left(\frac{x-3.758}{1.759}-\frac{y-1.199}{0.759}\right)^2+2(1-r)\frac{x-3.758}{1.759}\frac{y-1.199}{0.759}<0.583,$$

respectively.
- These regions contain more than the 50% and the 75% of the data (i.e. more than 75 and 113 data in this case).

Data sets. An example.

- Consider in the data set "*iris*" from R (Fisher, 1936), the variables $X = Petal.Length$ and $Y = Petal.Width$.
- We obtain $r = 0.9628654$ and R_1 and R_2 determined by

$$\left(\frac{x-3.758}{1.759}-\frac{y-1.199}{0.759}\right)^2+2(1-r)\frac{x-3.758}{1.759}\frac{y-1.199}{0.759}<0.292$$

and

$$\left(\frac{x-3.758}{1.759}-\frac{y-1.199}{0.759}\right)^2+2(1-r)\frac{x-3.758}{1.759}\frac{y-1.199}{0.759}<0.583,$$

respectively.
- These regions contain more than the 50% and the 75% of the data (i.e. more than 75 and 113 data in this case).

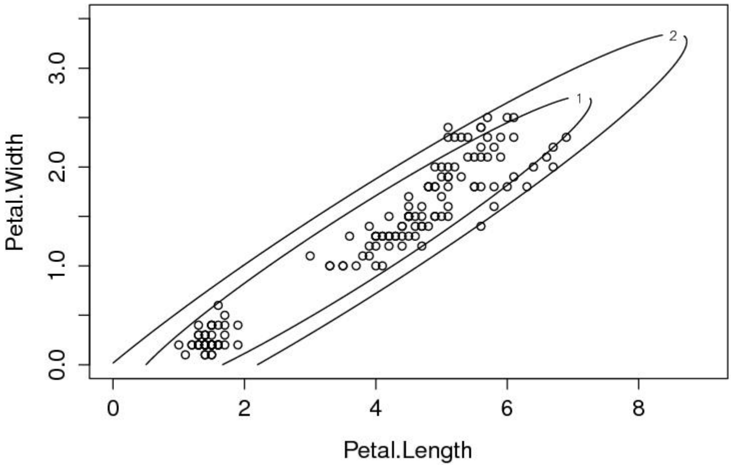

Figure: Regions R_1 and R_2 containing (for sure) at least the 50% and 75% of the data from $X = Petal.Length$ and $Y = Petal.Width$.

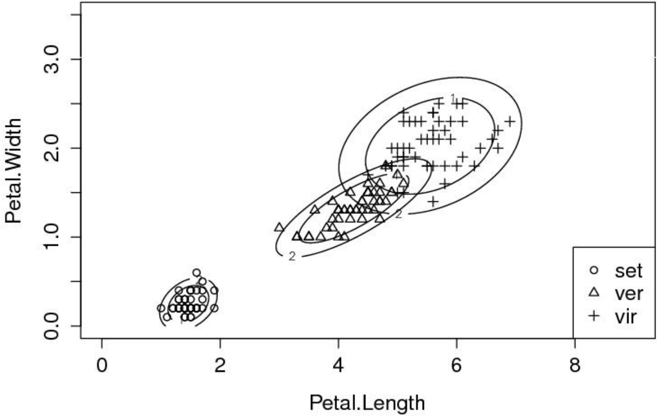

Figure: Regions R_1 and R_2 by species containing (for sure) at least the 50% and 75% of the data from $X = Petal.Length$ and $Y = Petal.Width$.

Data sets. An example.

- The two first principal components Y_1 and Y_2 of the four variables in this data set are

$$Y_1 = 0.521X_1^* - 0.269X_2^* + 0.580X_3^* + 0.565X_4^*$$

and

$$Y_2 = -0.377X_1^* - 0.923X_2^* - 0.025X_3^* - 0.067X_4^*,$$

where $X_i^* = (X_i - \overline{X}_i)/\sqrt{\widehat{V}_{i,i}}$, $i = 1, 2, 3, 4$.

- In this case, $\overline{Y}_1 = \overline{Y}_2 = 0$ and $r = 0$ and hence

$$R_1 = \{(x, y) : \frac{x^2}{2.918} + \frac{y^2}{0.914} < 4\}$$

and

$$R_2 = \{(x, y) : \frac{x^2}{2.918} + \frac{y^2}{0.914} < 8\}.$$

J. Navarro, E-mail: jorgenav@um.es

Data sets. An example.

- The two first principal components Y_1 and Y_2 of the four variables in this data set are

$$Y_1 = 0.521X_1^* - 0.269X_2^* + 0.580X_3^* + 0.565X_4^*$$

and

$$Y_2 = -0.377X_1^* - 0.923X_2^* - 0.025X_3^* - 0.067X_4^*,$$

where $X_i^* = (X_i - \overline{X}_i)/\sqrt{\widehat{V}_{i,i}}$, $i = 1, 2, 3, 4$.

- In this case, $\overline{Y}_1 = \overline{Y}_2 = 0$ and $r = 0$ and hence

$$R_1 = \{(x, y) : \frac{x^2}{2.918} + \frac{y^2}{0.914} < 4\}$$

and

$$R_2 = \{(x, y) : \frac{x^2}{2.918} + \frac{y^2}{0.914} < 8\}.$$

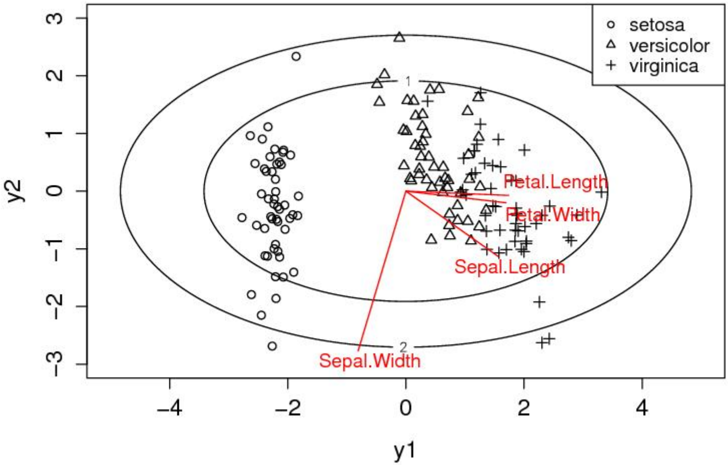

Figure: Regions R_1 and R_2 for the scores in the two first principal components containing (for sure) at least the 50% and 75% of the data scores.

References

- Chen, X. (2011). A new generalization of Chebyshev inequality for random vectors. arXiv:0707.0805v2.
- Navarro, J. (2013). A very simple proof of the multivariate Chebyshev's inequality. To appear in *Communications in Statistics – Theory and Methods*. Published online first Dec 2013. DOI:10.1080/03610926.2013.873135.
- Navarro, J. (2014). Can the bounds in the multivariate Chebyshev inequality be attained? *Statistics and Probability Letters* 91, 1–5.
- Navarro, J. (2014). A note on confidence regions based on the bivariate Chebyshev inequality. Applications to order statistics and data sets. *ISTATISTIK, Journal of the Turkish Statistical Association* 7, 1–14.

References

- Thank you for your attention!!