

Multivariate Distributions

Daniel B. Rowe, Ph.D.

Professor
Department of Mathematical and Statistical Sciences



Outline

Recall Univariate:
Uniforms to Normals
Normals to Chi-Square
Normal and Chi-Square to t

- **Bi(Multi)variate Normal Distribution**
- **Wishart Distribution (symmetric matrix)**
- **Bi(Multi)variate Student t**
- **Matrix Normal and Matrix T**

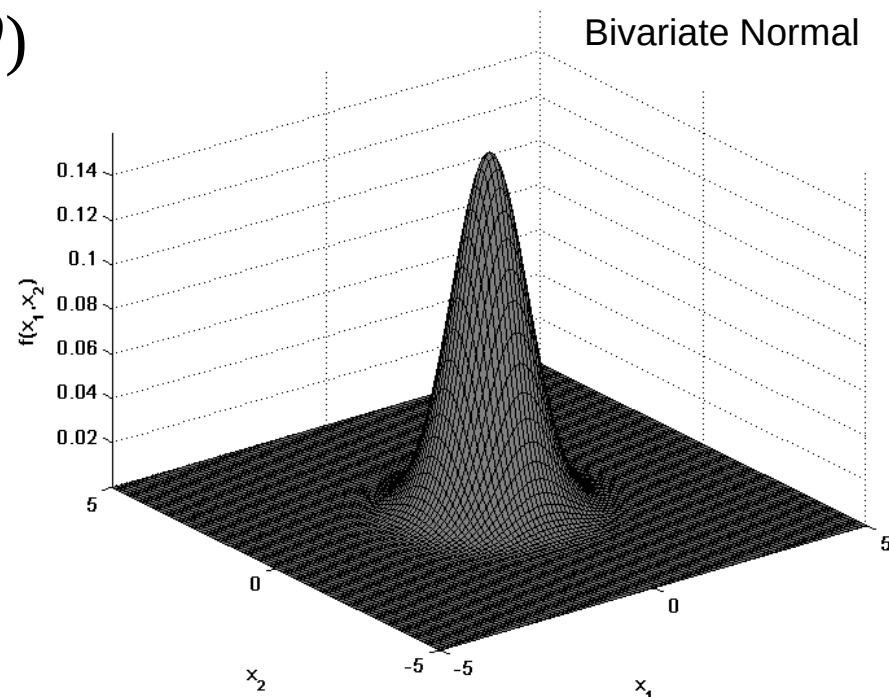
Bivariate Normal Distribution

A bivariate (2D) PDF $f(x_1, x_2 | \theta)$

of two continuous random variables (x_1, x_2) depending upon parameters θ satisfies

$$1) \quad 0 \leq f(x_1, x_2 | \theta), \quad \forall (x_1, x_2)$$

$$2) \quad \iint_{x_1 x_2} f(x_1, x_2 | \theta) dx_1 dx_2 = 1 \quad .$$



Bivariate Normal Distribution

Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be a 2-dimensional (or p -dimensional) random variable with PDF of x being $f(x | \theta)$, then

$$E(x | \theta) = \begin{pmatrix} E(x_1 | \theta) \\ E(x_2 | \theta) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

$= \mu_{2 \times 1}$

Marginal means. 

$$\text{var}(x | \theta) = \begin{pmatrix} \text{var}(x_1 | \theta) & \text{cov}(x_1, x_2 | \theta) \\ \text{cov}(x_1, x_2 | \theta) & \text{var}(x_2 | \theta) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

$= \Sigma_{2 \times 2}$

Marginal variances. 

Vectors are lower case, matrices are upper case.

Bivariate Normal Distribution

Recall: If $z \sim N(0,1)$, then

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} = (2\pi)^{-1/2} (1)^{-1/2} e^{-\frac{1}{2}(z-0)(1)^{-1}(z-0)}$$

↑ variance
↑ mean
↑ mean

Note the way I
have written this.



We can obtain a random variable x that has a

general normal distribution with mean μ and

variance σ^2 via the transformation $x = \sigma z + \mu$.

Bivariate Normal Distribution

The PDF of x can be obtained by

$$f(x | \mu, \sigma^2) = f(z(x)) \times |J(z \rightarrow x)| \quad J(z \rightarrow x) = \frac{dz(x)}{dx}$$

where $\underset{1 \times 1}{z} = z(x)$ and $J(\cdot)$ is the Jacobian of the transformation.

The PDF of $\underset{1 \times 1}{x}$ is

$$f(x | \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2}$$

Note the way I
have written this.



which can be written as

$$f(x | \mu, \sigma^2) = (2\pi)^{-1/2} (\sigma^2)^{-1/2} e^{-\frac{1}{2} (x - \mu)(\sigma^2)^{-1}(x - \mu)}$$

Bivariate Normal Distribution

Given two continuous random variables (z_1, z_2) , we write them as a 2-dimensional vector $\underset{2 \times 1}{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, and this vector

has the PDF $\underset{\text{vector}}{f_z}(z \mid \theta)$.
vector ↗

If z_1 and z_2 are independent, then

$$f_z(z \mid \theta) = f_{z_1}(z_1 \mid \theta_1) f_{z_2}(z_2 \mid \theta_2) .$$

Bivariate Normal Distribution

Let z_1 and z_2 be iid $N(0,1)$ random variables. Then, $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$

$$\begin{aligned} \text{has PDF } f_{z_1, z_2}(z_1, z_2) &= f_{z_1}(z_1) f_{z_2}(z_2) \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}(z_1^2 + z_2^2)}. \end{aligned}$$

With vector z , this can be rewritten as

$$f_z(z) = \frac{1}{2\pi} e^{-\frac{1}{2}z'z}.$$

Bivariate Normal Distribution

This can also be written as

$$f_Z(z) = \frac{1}{2\pi} e^{-\frac{1}{2}z'z}$$

$$= (2\pi)^{-2/2} |I_p|^{-1/2} e^{-\frac{1}{2}(z-0)'(I_p)^{-1}(z-0)}$$

$\nwarrow$ mean vector $\nwarrow$ mean vector
 $\nwarrow$ covariance matrix $\nwarrow$ covariance matrix

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

2×1

and we write that $z \sim N(0, I_2)$.

$\nwarrow$ mean vector $\nwarrow$ covariance matrix
 2×1

I_2 is 2×2
Identity matrix

That is, the 2-dimensional random vector z has a

2×1

mean vector of zero and identity variance-covariance matrix.

Bivariate Normal Distribution

This means that

$$f_Z(z) = (2\pi)^{-2/2} |I_p|^{-1/2} e^{-\frac{1}{2}(z-0)'(I_p)^{-1}(z-0)}$$

$$Z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

2×1

$$E(z) = \begin{pmatrix} E(z_1) \\ E(z_2) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$= \underset{2 \times 1}{\mu} = 0$$

2×2 Identity matrix

$$\text{var}(z) = \begin{pmatrix} \text{var}(z_1) & \text{cov}(z_1, z_2) \\ \text{cov}(z_1, z_2) & \text{var}(z_2) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \underset{2 \times 2}{\Sigma} = I_2$$

Bivariate Normal Distribution

If $\underset{2 \times 1}{z} \sim N(0, I_2)$, then

$$f(z) = (2\pi)^{-2/2} |I_p|^{-1/2} e^{-\frac{1}{2}(z-0)'(I_p)^{-1}(z-0)}$$

We can obtain a random variable $\underset{2 \times 1}{x}$ that has a general

normal distribution with mean vector $\underset{2 \times 1}{\mu}$ and variance-

covariance matrix $\underset{2 \times 2}{\Sigma}$ via the transformation $\underset{2 \times 1}{x} = \underset{2 \times 2}{A} \underset{2 \times 1}{z} + \underset{2 \times 1}{\mu}$

where $\underset{2 \times 2}{\Sigma} = \underset{2 \times 2}{A} \underset{2 \times 2}{A}'$, is a factorization (i.e. Cholesky or Eigen).

Bivariate Normal Distribution

If a random variable x has a normal distribution with

2×1

mean vector μ and variance-covariance matrix Σ , then

2×1

2×2

$$f(x | \mu, \Sigma) = (2\pi)^{-p/2} |\Sigma|^{-1/2} e^{-\frac{1}{2}(x-\mu)' \Sigma^{-1} (x-\mu)}$$

Annotations for the PDF equation:

- μ : mean vector
- Σ : covariance matrix
- Σ^{-1} : covariance matrix
- $(x-\mu)' \Sigma^{-1} (x-\mu)$: mean vector

$$x, \mu \in \mathbb{R}^p$$

$$p = 2$$

$$\Sigma > 0$$

↑ set of pos
def matrices

and we write $x \sim N(\mu, \Sigma)$. The covariance matrix Σ , has to

2×1

2×1

2×2

2×2

be of full rank (there is an inverse in PDF).

↑ make sure you know what this means

Bivariate Normal Distribution

Let's take a closer look at this bivariate transformation.

$$\underset{2 \times 1}{x} = \underset{2 \times 2}{A} \underset{2 \times 1}{z} + \underset{2 \times 1}{\mu}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

$$x_1 = a_{11}z_1 + a_{12}z_2 + \mu_1$$

$$x_2 = a_{21}z_1 + a_{22}z_2 + \mu_2$$

We can solve for $\underset{1 \times 1}{z_1}$ and $\underset{1 \times 1}{z_2}$ in terms of $\underset{1 \times 1}{x_1}$ and $\underset{1 \times 1}{x_2}$.

Bivariate Normal Distribution

This will give us

$$\underset{2 \times 1}{x} = \underset{2 \times 2}{A} \underset{2 \times 1}{z} + \underset{2 \times 1}{\mu}$$

$$\underset{2 \times 1}{z} = \underset{2 \times 2}{A}^{-1} (\underset{2 \times 1}{x} - \underset{2 \times 1}{\mu})$$

A invertible

$$\underset{2 \times 2}{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

$$\leftarrow \underset{2 \times 2}{B} = \underset{2 \times 2}{A}^{-1} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$\underset{2 \times 1}{\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}} = \underset{2 \times 2}{\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}} \underset{2 \times 1}{\begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix}}$$

$$z_1 = b_{11}(x_1 - \mu_1) + b_{12}(x_2 - \mu_2)$$

$$z_2 = b_{21}(x_1 - \mu_1) + b_{22}(x_2 - \mu_2)$$

Bivariate Normal Distribution

$$z_1 = b_{11}(x_1 - \mu_1) + b_{12}(x_2 - \mu_2)$$

$$z_2 = b_{21}(x_1 - \mu_1) + b_{22}(x_2 - \mu_2)$$

Continuing on, this leads to

$$J(z_1, z_2 \rightarrow x_1, x_2) = \begin{vmatrix} \frac{\partial z_1(x_1, x_2)}{\partial x_1} & \frac{\partial z_1(x_1, x_2)}{\partial x_2} \\ \frac{\partial z_2(x_1, x_2)}{\partial x_1} & \frac{\partial z_2(x_1, x_2)}{\partial x_2} \end{vmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = B_{2 \times 2}$$

i.e. with $z = B(x - \mu)$, the vector derivative is $J = \frac{\partial z}{\partial x} = B_{2 \times 2}$

Bivariate Normal Distribution

$$z = B(x - \mu)$$

The distribution of vector variable x (joint of x_1 and x_2) is

 2×1

$$f_X(x | \theta) = f_Z(z(x)) \times |J(z \rightarrow x)|$$

$$f(z) = (2\pi)^{-2/2} |I_n|^{-1/2} e^{-\frac{1}{2}(z-0)'(I_n)^{-1}(z-0)} \quad J = \frac{\partial z}{\partial x} = B \quad 2 \times 2$$

$$f_X(x | \mu, \Sigma) = (2\pi)^{-2/2} |I_n|^{-1/2} e^{-\frac{1}{2}(B(x-\mu)-0)'(I_n)^{-1}(B(x-\mu)-0)} |B|$$

$$\Sigma = AA' \quad , \quad |\Sigma| = |A| |A'| = |A|^2 \quad , \quad |\Sigma|^{1/2} = |A| \quad , \quad B = A^{-1} \quad , \quad |\Sigma|^{-1/2} = |B|$$

$$f_X(x | \mu, \Sigma) = (2\pi)^{-2/2} |\Sigma|^{-1/2} e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)} \quad \begin{array}{l} x, \mu \in \mathbb{R}^p \\ \Sigma > 0 \end{array}$$

Bivariate Normal Distribution

This form may be more familiar

$$f_X(x_1, x_2 | \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q}$$

$$Q = \frac{1}{(1-\rho^2)} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right]$$

$$\sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$$

$$\rho = \sigma_{12} / (\sigma_1 \sigma_2)$$

$$\sigma_{12} = \text{COV}(x_1, x_2)$$

Bivariate Normal Distribution

Theorem:

If x is a 2-D (or p -D) random variable from $f(x|\mu, \Sigma)$, with

$$E(x | \mu, \Sigma) = \underset{p \times 1}{\mu} \quad \text{think of } p=2$$

$$\text{var}(x | \mu, \Sigma) = \underset{p \times p}{\Sigma}$$

then we form $y = \underset{r \times 1}{A} \underset{r \times p}{x} + \underset{r \times 1}{\delta}$ where dimensions match

and $\underset{r \times p}{A}$ full column rank ($A: r \times p, r \leq p$), then

$$E(y | \mu, \Sigma, \delta, A) = \underset{r \times p}{A} \underset{p \times 1}{\mu} + \underset{r \times 1}{\delta}$$

$$\text{var}(y | \mu, \Sigma, A) = \underset{r \times p}{A} \underset{p \times p}{\Sigma} \underset{r \times p}{A'}$$

Bivariate Normal Distribution

Recall: Let $u_1 \sim \text{uniform}(0,1)$ and $u_2 \sim \text{uniform}(0,1)$.

$$\text{Let } z_1 = \sqrt{-2\ln(u_1)} \cos(2\pi u_2) \text{ and } z_2 = \sqrt{-2\ln(u_1)} \sin(2\pi u_2) \\ \text{then } u_1(z_1, z_2) = e^{-\frac{1}{2}(z_1^2 + z_2^2)} \text{ and } u_2(z_1, z_2) = \frac{1}{2\pi} \text{atan}\left(\frac{z_2}{z_1}\right).$$

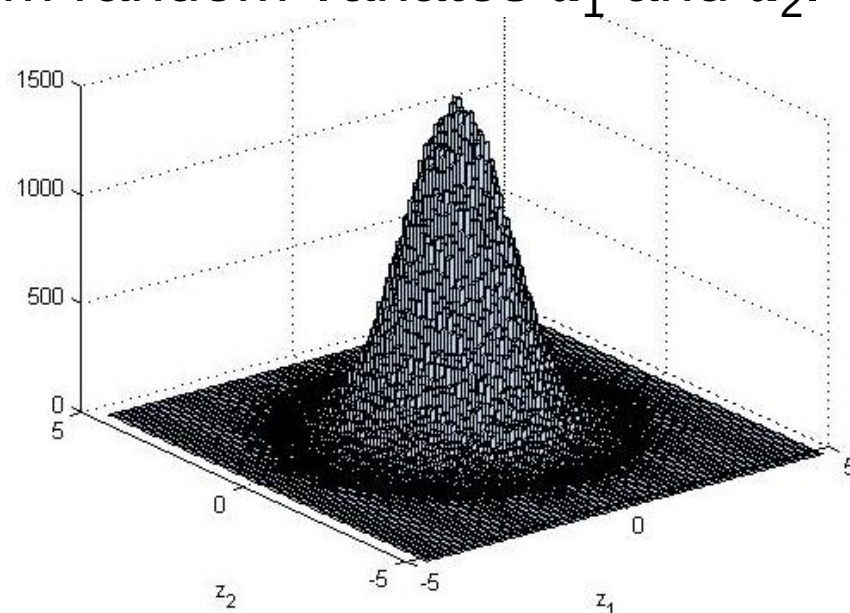
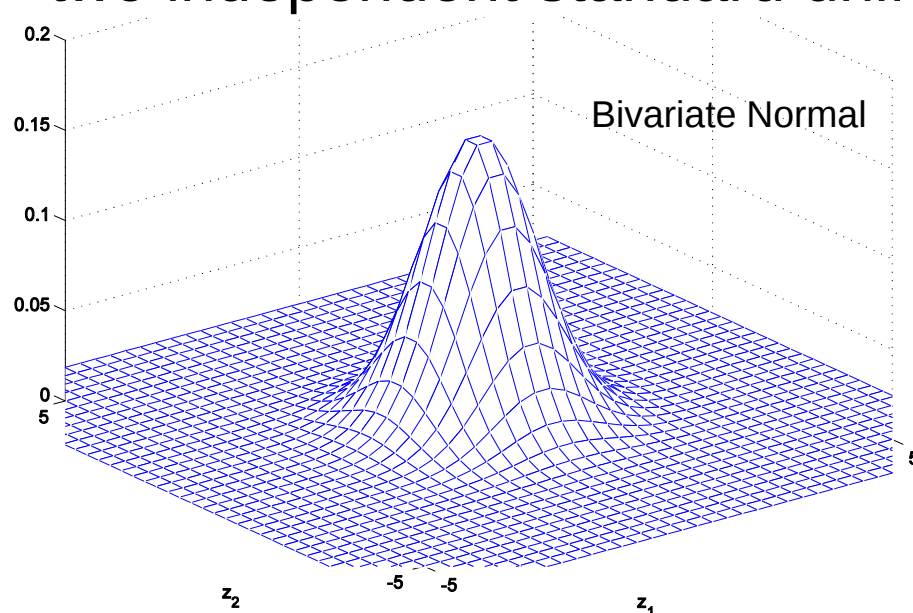
$$f_{z_1, z_2}(z_1, z_2 | \theta) = f_{u_1, u_2}(u_1(z_1, z_2), u_2(z_1, z_2) | \theta) \times |J(u_1, u_2 \rightarrow z_1, z_2)|$$

$$f_{z_1, z_2}(z_1, z_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_2^2}$$

This means $z_1 \sim N(0,1)$, $z_2 \sim N(0,1)$, z_1 and z_2 are independent.

Bivariate Normal Distribution

Obtain 2-D standard normal variates $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}_{2 \times 1}$ by transforming two independent standard uniform random variates u_1 and u_2 .



First half of 10^6 standard normal variates as z_1 's and second half as z_2 's. Produce 5×10^5 statistically independent z 's.

Bivariate Normal Distribution

Multiplied 5×10^5 simulated z 's by A and added μ .
 The y 's are now $N(\mu, \Sigma = AA')$.

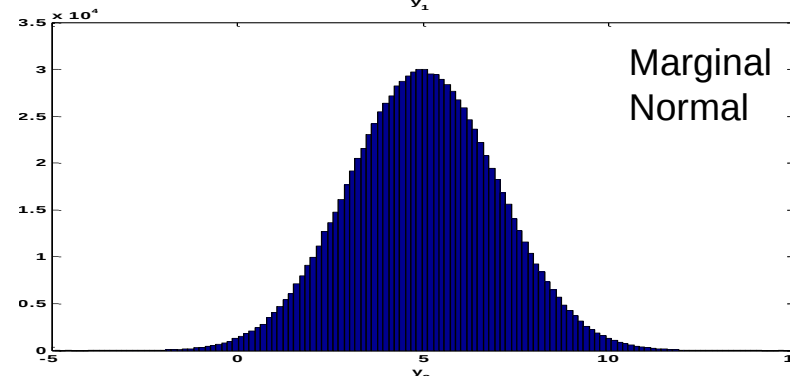
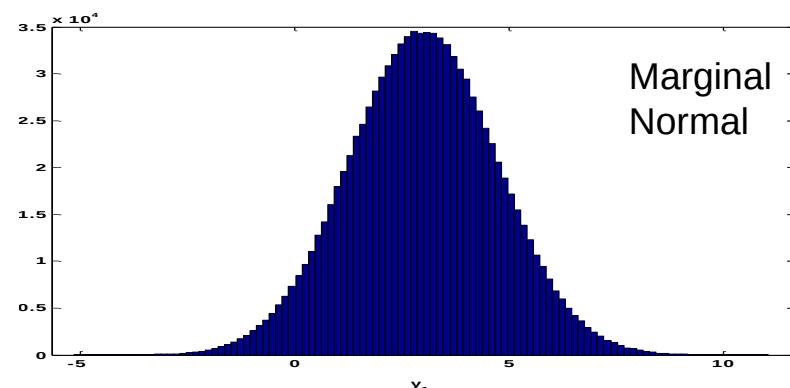
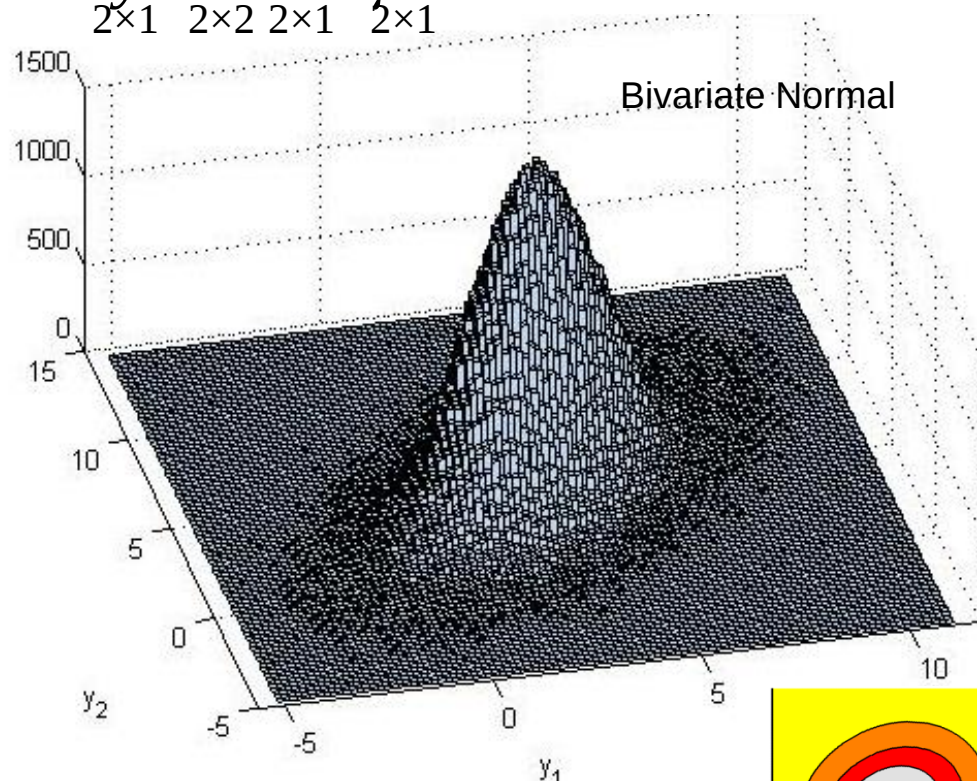
$$y = Az + \mu$$

$\begin{matrix} 2 \times 1 & 2 \times 2 & 2 \times 1 & 2 \times 1 \end{matrix}$

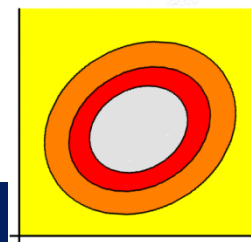
$$\mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix} \quad \rho = 0.58$$

Cholesky

$$A = \begin{pmatrix} 1.7321 & 0 \\ 1.1547 & 1.6330 \end{pmatrix}$$



cross
section



$$A = \text{chol}(\Sigma)$$

Bivariate Normal Distribution

Theorem:

If x_1 and x_2 are independent 2-D (or p -D) RVs with

$$E(x_1 | \mu_1, \Sigma_1) = \mu_1 \quad E(x_2 | \mu_2, \Sigma_2) = \mu_2 \quad \text{think of } p=2$$

$p \times 1 \qquad p \times 1 \qquad p \times 1 \qquad p \times 1$

$$\text{var}(x_1 | \mu_1, \Sigma_1) = \Sigma_1 \quad \text{var}(x_2 | \mu_2, \Sigma_2) = \Sigma_2$$

$p \times p \qquad p \times p \qquad p \times p \qquad p \times p$

then if we let $y = (x_1 + x_2) / 2$,

$p \times 1 \qquad p \times 1$

$$E(y | \mu_1, \Sigma_1, \mu_2, \Sigma_2) = (\mu_1 + \mu_2) / 2$$

$p \times 1 \qquad p \times 1$

$$\text{var}(y | \mu_1, \Sigma_1, \mu_2, \Sigma_2) = (\Sigma_1 + \Sigma_2) / 4$$

$p \times p \qquad p \times p$

Recall in 1D if we let

$$\mu = \mu_1 = \mu_2 \quad \text{and} \\ \sigma^2 = \sigma_1^2 = \sigma_2^2 \quad \text{then} \\ E(y | \mu) = \mu \quad \text{and} \\ \text{var}(y | \mu, \sigma^2) = \sigma^2 / 2 .$$

Bivariate Normal Distribution

Theorem:

If $x_1, \dots, x_n$ are independent 2-D (or p -D) RVs with

$$E(x_i \mid \mu, \Sigma) = \mu_{p \times 1} \quad i = 1, \dots, n \quad \text{think of } p=2$$

$$\text{var}(x_i \mid \mu, \Sigma) = \Sigma_{p \times p}$$

then if we let $\bar{x} = (x_1 + \dots + x_n) / n$,
 $\bar{x}_{p \times 1}$

$$E(\bar{x} \mid \mu, \Sigma) = \mu_{p \times 1}$$

$$\text{var}(\bar{x} \mid \mu, \Sigma) = \Sigma_{p \times p} / n \quad .$$

If x 's are normal, then

$$\bar{x}_{p \times 1} \sim N(\mu_{p \times 1}, \Sigma_{p \times p} / n)$$

Otherwise CLT type result?

Bivariate Normal Distribution

If $x_1, \dots, x_n$ are IID $N(\mu, \Sigma)$ 2-D (or p -D) RVs and

$$\bar{x} = (x_1 + \dots + x_n) / n \sim N(\mu, \Sigma / n)$$

then

$$f_{\bar{x}}(\bar{x} | \mu, \Sigma) = (2\pi)^{-2/2} |\Sigma / n|^{-1/2} e^{-\frac{1}{2}(\bar{x} - \mu)'(\Sigma/n)^{-1}(\bar{x} - \mu)}.$$

Also note that

$$n^{1/2} \Sigma^{-1/2} (\bar{x} - \mu) \sim N(0, I_2) \quad \leftarrow \begin{array}{l} \text{Multivariate version of} \\ \frac{(\bar{x} - \mu)}{\sigma / \sqrt{n}} \sim N(0, 1) \end{array}$$

$$\Sigma^{-1/2}_{p \times p} = A^{-1}_{p \times p} = B_{p \times p}$$

Estimate Σ by its MLE $\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})'$.

Wishart Distribution

Multivariate version of gamma distribution.

A random $p \times p$ matrix variate G follows the Wishart

distribution with scale matrix Σ and ν df denoted $G \sim W(\Sigma, \nu)$

$$\text{iff } f(G | \Sigma, \nu) = k_W |\Sigma|^{-\frac{\nu}{2}} |G|^{-\frac{\nu-p-1}{2}} e^{-\frac{1}{2} \text{tr}(\Sigma^{-1}G)} \quad \begin{array}{l} G, \Sigma > 0 \\ \nu > p+1 \\ \nu \in \mathbb{N} \end{array}$$

$$\text{where } k_W^{-1} = 2^{\frac{\nu p}{2}} \pi^{-\frac{p(p-1)}{4}} \prod_{j=1}^p \Gamma\left(\frac{\nu+1-j}{2}\right)$$

$$\text{If } p=1, f(g | \sigma^2, \nu) = \frac{g^{\frac{\nu}{2}-1} e^{-\frac{g}{2\sigma^2}}}{\Gamma(\nu/2)(2\sigma^2)^{\nu/2}}$$

Gamma dist by $\alpha = \nu / 2$, Chi-square by $y = g / \sigma^2$.

Wishart Distribution

$$f(G | \Sigma, \nu) = k_W |\Sigma|^{-\frac{\nu}{2}} |G|^{-\frac{\nu-p-1}{2}} e^{-\frac{1}{2} \text{tr}(\Sigma^{-1}G)}$$

$$k_W^{-1} = 2^{\frac{\nu p}{2}} \pi^{-\frac{p(p-1)}{4}} \prod_{j=1}^p \Gamma\left(\frac{\nu+1-j}{2}\right)$$

The mean, variance, and covariance of elements are

$$E(G | \Sigma, \nu) = \nu \Sigma_{p \times p}$$

$$\text{var}(G_{ij} | \Sigma, \nu) = \nu(\Sigma_{ij}^2 + \Sigma_{ii}\Sigma_{jj})$$

$$\text{cov}(G_{ij}G_{kl} | \Sigma, \nu) = \nu(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk})$$

Wishart Distribution

We obtain a (singular) $W(\Sigma, 1)$ variate G_1 by transforming a centered normal vector variate $z_1 = (x_1 - \mu)$ via $G_1 = z_1 z_1'$ and a $W(\Sigma, n)$ variate by $G = \sum_{i=1}^n z_i z_i'$. If the mean vector is not known, then we can estimate it and lose one df .

$$\underbrace{\sum_{i=1}^n (x_i - \mu)'(x_i - \mu)}_{W(\Sigma, n), n \geq p} \overset{G}{=} \underbrace{\sum_{i=1}^n (x_i - \bar{x})'(x_i - \bar{x})}_{W(\Sigma, n-1), n-1 > p} \overset{G_2}{+} \underbrace{n(\bar{x} - \mu)'(\bar{x} - \mu)}_{W(\Sigma, 1), \text{ singular}} \overset{G_1}{}$$

add and subtract $\bar{x}$ in parentheses

Let $p=1$ and get σ^2 times χ^2 result.

Wishart Distribution

Theorem:

If G is a $p \times p$ random matrix variable from $f(G|\Sigma, \nu)$, with

$$f(G|\Sigma, \nu) = k_W |\Sigma|^{-\frac{\nu}{2}} |G|^{-\frac{\nu-p-1}{2}} e^{-\frac{1}{2}tr(\Sigma^{-1}G)}$$

then if we form $Q = AGA'$ where dimensions match

and A full row rank ($A: r \times p, r \leq p$), then $Q \sim W(\Delta = A\Sigma A', \nu)$

$$E(Q|\Delta, \nu) = \nu\Delta$$

$$\text{var}(Q_{ij}|\Delta, \nu) = \nu(\Delta_{ij}^2 + \Delta_{ii}\Delta_{jj})$$

$$\text{cov}(Q_{ij}Q_{kl}|\Delta, \nu) = \nu(\Delta_{ik}\Delta_{jl} + \Delta_{il}\Delta_{jk}) .$$

Wishart Distribution

$$\mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix}_{2 \times 1} \quad \Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}_{2 \times 2}$$

Took 5×10^4 sets of $n=10$ variates x , subtracted mean $\mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix}_{2 \times 1}$

from each set, transpose multiplied each value, added the

10 values in set to form each G . The G 's are now $W(\Sigma, \nu = n)$.

$$E(G \mid \Sigma, \nu) = \nu \Sigma$$

$$\nu = 10 \quad \Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$$

$$\text{var}(G_{ij} \mid \Sigma, \nu) = \nu(\Sigma_{ij}^2 + \Sigma_{ii}\Sigma_{jj})$$

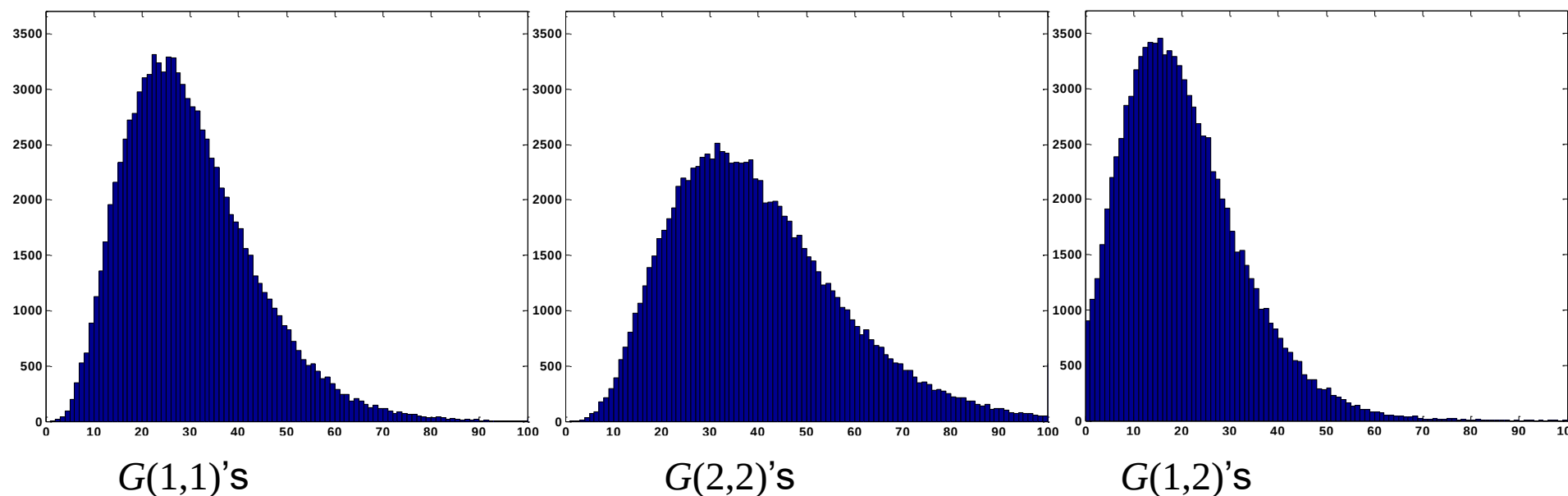
$$\text{cov}(G_{ij}G_{kl} \mid \Sigma, \nu) = \nu(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk})$$

$$G = \sum_{i=1}^n (x_i - \mu)'(x_i - \mu)$$

Wishart Distribution

$$\mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix}_{2 \times 1} \quad \Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}_{2 \times 2}$$

The G 's, $G = \sum_{i=1}^n (x_i - \mu)'(x_i - \mu)$ are now $W(\Sigma, \nu)$. $\nu = 10$



$$E(G \mid \Sigma, \nu) = \nu \Sigma = \begin{pmatrix} 30 & 20 \\ 20 & 40 \end{pmatrix} \quad \text{var}(G_{ij} \mid \Sigma, \nu) = \nu(\Sigma_{ij}^2 + \Sigma_{ii}\Sigma_{jj}) = \begin{pmatrix} 180 & 160 \\ 160 & 320 \end{pmatrix}$$

$$\text{cov}(G_{ij}G_{kl} \mid \Sigma, \nu) = \nu(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk}) = \begin{matrix} 80 & 120 & 160 \\ 11,22 & 11,12 & 22,12 \end{matrix} \leftarrow ij,kl$$

Wishart Distribution

Took 5×10^4 sets of $n=10$ variates x , subtracted mean $\bar{x}$ from each set, transpose multiplied each value, added the 10 values to form each G_2 . The G_2 's are now $W(\Sigma, \nu = n - 1)$.

$$E(G_2 \mid \Sigma, \nu) = \nu \Sigma \qquad \nu = 9 \quad \Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$$

$$\text{var}(G_{2ij} \mid \Sigma, \nu) = \nu(\Sigma_{ij}^2 + \Sigma_{ii}\Sigma_{jj})$$

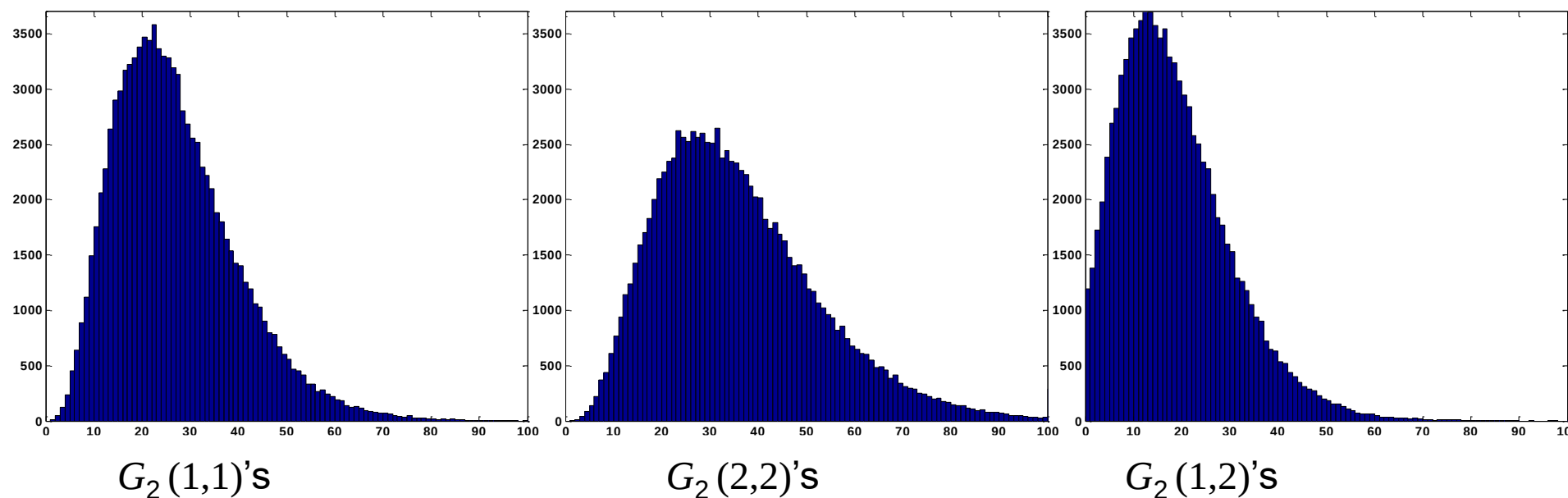
$$\text{cov}(G_{2ij}G_{2kl} \mid \Sigma, \nu) = \nu(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk})$$

Wishart Distribution

$$\mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix}_{2 \times 1} \quad \Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}_{2 \times 2}$$

Toggle with
Next slide

The G_2 's, $G_2 = \sum_{i=1}^n (x_i - \bar{x})'(x_i - \bar{x})$ are now $W(\Sigma, \nu)$. $\nu = 9$



$$E(G_2 | \Sigma, \nu) = \nu \Sigma = \begin{pmatrix} 27 & 18 \\ 18 & 36 \end{pmatrix} \quad \text{var}(G_{2ij} | \Sigma, \nu) = \nu(\Sigma_{ij}^2 + \Sigma_{ii}\Sigma_{jj}) = \begin{pmatrix} 162 & 144 \\ 144 & 288 \end{pmatrix}$$

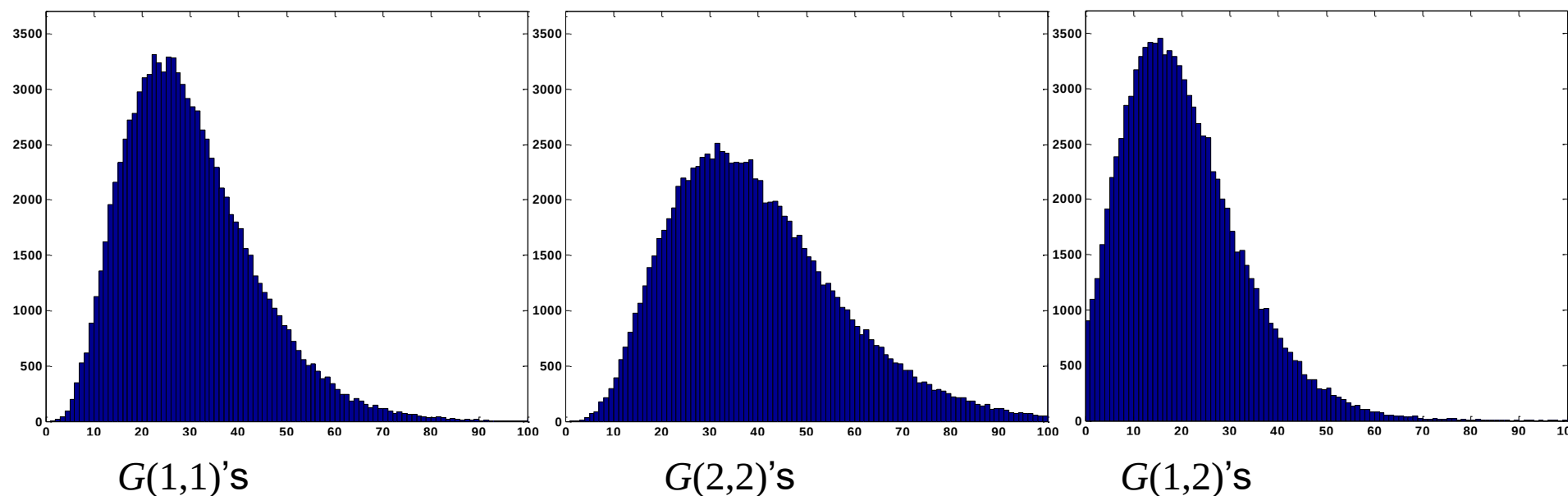
$$\text{cov}(G_{2ij} G_{2kl} | \Sigma, \nu) = \nu(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk}) = \begin{matrix} 72 & 108 & 144 \\ 11,22 & 11,12 & 22,12 \end{matrix} \leftarrow ij,kl$$

Wishart Distribution

$$\mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix}_{2 \times 1} \quad \Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}_{2 \times 2}$$

Toggle with
Previous slide

The G 's, $G = \sum_{i=1}^n (x_i - \mu)'(x_i - \mu)$ are now $W(\Sigma, \nu)$. $\nu = 10$



$$E(G \mid \Sigma, \nu) = \nu \Sigma = \begin{pmatrix} 30 & 20 \\ 20 & 40 \end{pmatrix} \quad \text{var}(G_{ij} \mid \Sigma, \nu) = \nu(\Sigma_{ij}^2 + \Sigma_{ii}\Sigma_{jj}) = \begin{pmatrix} 180 & 160 \\ 160 & 320 \end{pmatrix}$$

$$\text{cov}(G_{ij}G_{kl} \mid \Sigma, \nu) = \nu(\Sigma_{ik}\Sigma_{jl} + \Sigma_{il}\Sigma_{jk}) = \begin{matrix} 80 & 120 & 160 \\ 11,22 & 11,12 & 22,12 \end{matrix} \leftarrow ij,kl$$

Multivariate Student-t Distribution

A 2-D ($p=2$) random vector variate t follows a standard $p \times 1$

Student t distribution with ν *df* denoted $t \sim t(\nu)$ iff $p \times 1$

$$f(t | \nu) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{(\nu\pi)^{\frac{p}{2}} \Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\left[1 + \frac{1}{\nu} t' t\right]^{\frac{\nu+p}{2}}}, \text{ where } t \in \mathbb{R}^p, \nu \in \mathbb{N}.$$

The mean and variance of t are

$$E(t | \nu) = 0_{p \times 1}$$

$$\nu > 1$$

$$\text{var}(t | \nu) = \frac{\nu}{\nu - 2} I_p$$

$$\nu > 2$$

Multivariate Student-t Distribution

We obtain a standard $t(\nu=n-1)$ variate t by transforming a normal 2-D (p -D) random variate $\bar{x}_{p \times 1}$ and a Wishart random matrix variate G that are independent

$$\begin{pmatrix} t \\ V \end{pmatrix} = \begin{pmatrix} n^{1/2} \nu^{1/2} G_2^{-1/2} (\bar{x} - \mu) \\ G_2 \end{pmatrix} \quad \bar{x}_{p \times 1} \sim N(\mu, \Sigma / n) \quad G_2 \sim W(\Sigma, \nu)_{p \times p} .$$

The original variables in terms of the new variables

$$\begin{pmatrix} \bar{x}(t, \nu) \\ G_2(t, V) \end{pmatrix} = \begin{pmatrix} n^{-1/2} \nu^{-1/2} V^{1/2} t + \mu \\ V \end{pmatrix}$$

Multivariate Student-t Distribution

The joint distribution of (t, V) can be obtained as

$$f(t, V | \mu, \Sigma, \nu) = f(\bar{x}(t, V), G_2(t, V)) \times |J(\bar{x}, G_2 \rightarrow t, V)|$$

The Jacobian of the transformation is

$$|J(\bar{x}, G_2 \rightarrow t, V)| = n^{-1/2} \nu^{-1/2} V^{1/2}$$

The joint PDF of (t, V) is

$$p(t, V | \Sigma, \nu) = (2\pi\nu)^{-p/2} k_W |\Sigma|^{-\frac{\nu+1}{2}} |V|^{-\frac{(\nu+1)-p-1}{2}} e^{-\frac{1}{2} \text{tr} \Sigma^{-1} V (I_p + \frac{1}{\nu} t t')}$$

The distribution of t can be found by integrating out V

$$f(t | \nu) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{(\nu\pi)^{\frac{p}{2}} \Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\left[1 + \frac{1}{\nu} t' t\right]} \cdot$$

After integration use

$$|I_p + \frac{1}{\nu} t t'| = \left|1 + \frac{1}{\nu} t' t\right|$$

Multivariate Student-t Distribution

Recall $\bar{x} \sim N(\mu, \Sigma / n)$

so $n^{1/2} \Sigma^{-1/2} (\bar{x} - \mu) \sim N(0, I)$

also $\sum_{i=1}^n (x_i - \bar{x})' (x_i - \bar{x}) \sim W(\Sigma, n-1)$

thus $(n-1) \Sigma^{-1/2} \left[\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})' (x_i - \bar{x}) \right] (\Sigma^{-1/2})' \sim W(I, n-1)$

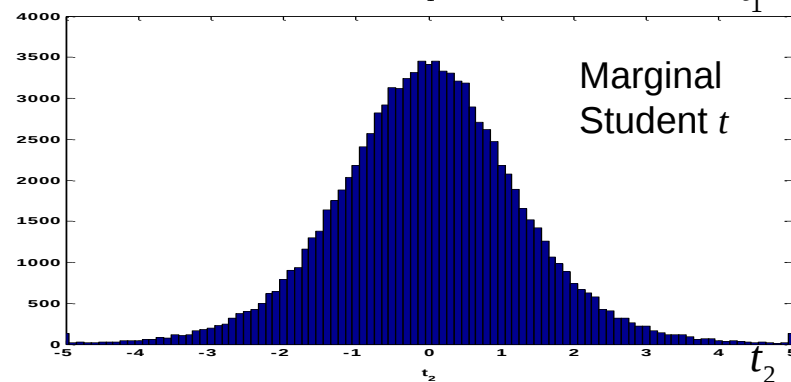
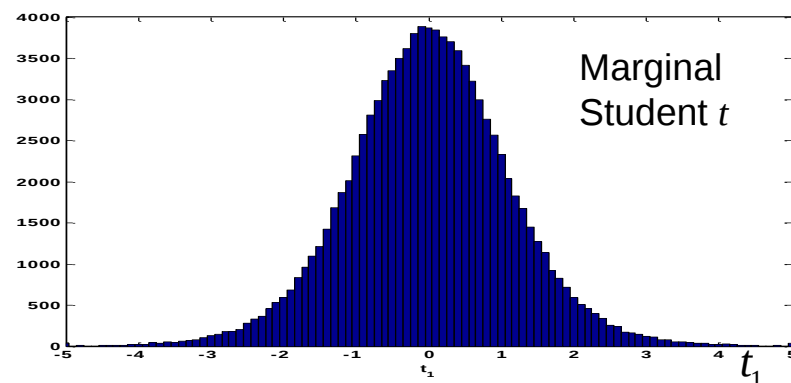
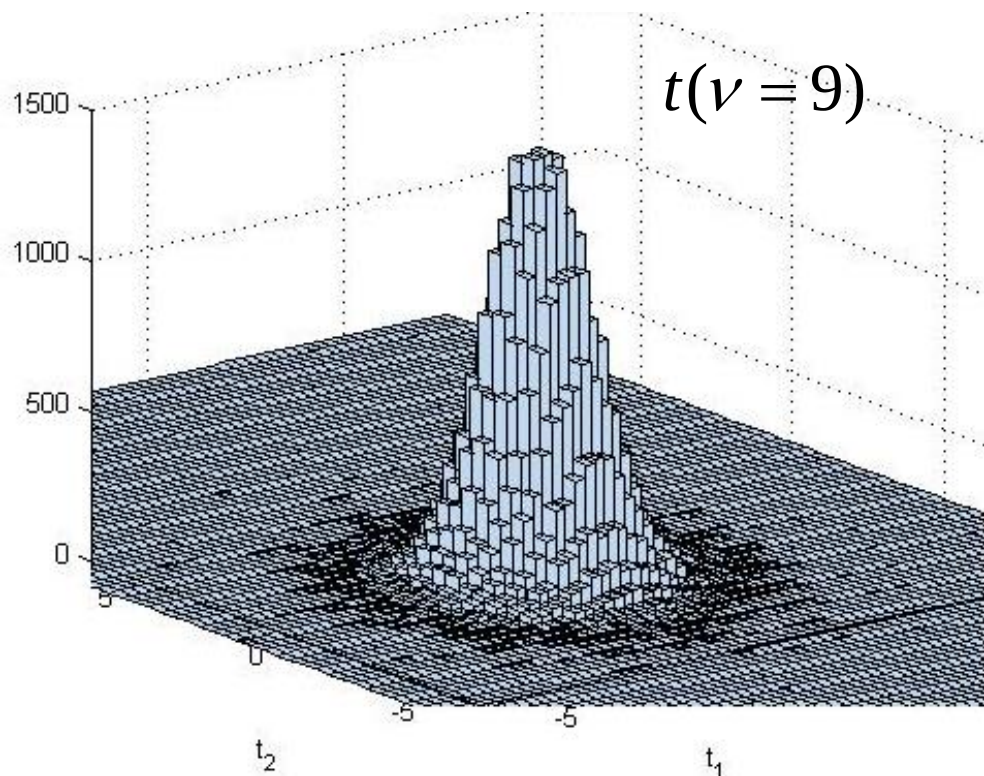
but $\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})' (x_i - \bar{x}) = S$

finally $[(n-1)S]^{-1/2} n^{1/2} (\bar{x} - \mu) \sim t(n-1)$

Multivariate Student-t Distribution

$$\mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$$

Take sample mean $\bar{x}$ from each 5×10^4 sets of 10 variates, subtracted μ from each set, premultiplied the 5×10^4 centered normal variates by the $5 \times 10^4 [(n-1)S]^{-1/2} n^{1/2} (\bar{x} - \mu)$



Multivariate Student-t Distribution

A 2-D ($p=2$) random vector variate s follows a general $p \times 1$

Student t distribution with location μ and scale Σ and $df \ \nu$
 $s \sim t(\nu, \mu, \Sigma)$ $p \times 1$ $p \times p$

$$f(s | \nu, \mu, \Sigma) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{(\nu\pi)^{\frac{p}{2}} \Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\left[1 + \frac{1}{\nu}(s - \mu)' \Sigma^{-1} (s - \mu)\right]^{\frac{\nu+p}{2}}}$$

where $s, \mu \in \mathbb{R}^p$, $\nu \in \mathbb{N}$, $\Sigma > 0$.

The mean and variance of s are

$$E(t | \nu, \mu, \Sigma) = \mu_{p \times 1} \quad \nu > 1 \quad \text{var}(s | \nu, \mu, \Sigma) = \frac{\nu}{\nu - 2} \Sigma_{p \times p} \quad \nu > 2$$

Multivariate Distributions

We can also perform hypothesis tests.

Single mean $H_o : \underset{p \times 1}{\mu} = \underset{p \times 1}{\mu_0}$ vs $H_1 : \underset{p \times 1}{\mu} \neq \underset{p \times 1}{\mu_0}$

Two Means $H_o : \underset{p \times 1}{\mu_1} = \underset{p \times 1}{\mu_2}$ vs $H_1 : \underset{p \times 1}{\mu_1} \neq \underset{p \times 1}{\mu_2}$

MANOVA $H_o : \underset{p \times 1}{\mu_1} = \underset{p \times 1}{\mu_2} = \underset{p \times 1}{\mu_3}$ vs $H_1 : \underset{p \times 1}{\mu}'\text{s not all equal}$

Multivariate Regression $Y = X B + E$ $\hat{B} = (X'X)^{-1}X'Y$
 $\underset{n \times p}{Y} = \underset{n \times (q+1)}{X} \underset{(q+1) \times p}{B} + \underset{n \times p}{E}$
 $\underset{p \times p}{S} = \frac{1}{n-q-1} (Y - X\hat{B})'(Y - X\hat{B})$

$H_o : \underset{p \times p}{\Sigma} = \text{diagonal}$ vs $H_1 : \underset{p \times p}{\Sigma} \neq \text{diagonal}$

Matrix Distributions

Matrix Normal and Matrix T distributions also exist.

$$f(X | M, \Sigma, \Phi) = (2\pi)^{-np/2} |\Phi|^{-p/2} |\Sigma|^{-n/2} e^{-\frac{1}{2} \text{tr} \Phi^{-1} (X-M)' \Sigma^{-1} (X-M)}$$

$p \times n$

$$f(T | \nu, T_0, \Sigma, \Phi) = k_T \frac{|\Phi|^{\frac{\nu}{2}} |\Sigma|^{-\frac{n}{2}}}{\left[\Phi + \frac{1}{\nu} (T - T_0)' \Sigma^{-1} (T - T_0) \right]^{\frac{\nu+p}{2}}}$$

$p \times n$

$$k_T = \frac{\prod_{j=1}^n \Gamma\left(\frac{\nu+p+1-j}{2}\right)}{(\nu\pi)^{\frac{np}{2}} \prod_{j=1}^n \Gamma\left(\frac{\nu+1-j}{2}\right)}$$

$n=1 \rightarrow$ multivariate
 $n, p=1 \rightarrow$ univariate

Homework 12:

- 1) Repeat the simulation processes performed in this lecture.
Generate two uniforms then transform to two standard normals then to two normals $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ with $\mu = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$ using the Cholesky factorization. $A = \text{chol}(\Sigma)'$;
Repeat 5 times to get a sample of size $n=5$.
Repeat 10^4 times to get 10^4 bivariate samples of $n = 5$.
 - a) Compute bivariate mean $\bar{x}$ for each sample.
 - b) Compute bivariate covariance matrix S for each sample.
 - c) For each sample compute $[(n-1)S]^{-1/2} n^{1/2} (\bar{x} - \mu)$.
 - d) Make histograms of items (elements) in a), b), and c).

Homework 12:

- 2) Repeat simulation process with eigenvalue-eigenvector decomposition $\Sigma = UD^{1/2}(D^{1/2}U)'$ instead of Cholesky $\Sigma = AA'$.
[v,d]=eig(Sigma); A=v*sqrt(d);
- 3) Comment of similarities and/or differences between 1) and 2).