Bivariate Probability Density Functions

Dr. Daniel B. Rowe
Professor of Computational Statistics
Department of Mathematical and Statistical Sciences
Marquette University

Outline

Motivation

Properties and Summaries

Bivariate Normal PDF

Generating Bivariate Normals

Discussion

Homework

Motivation

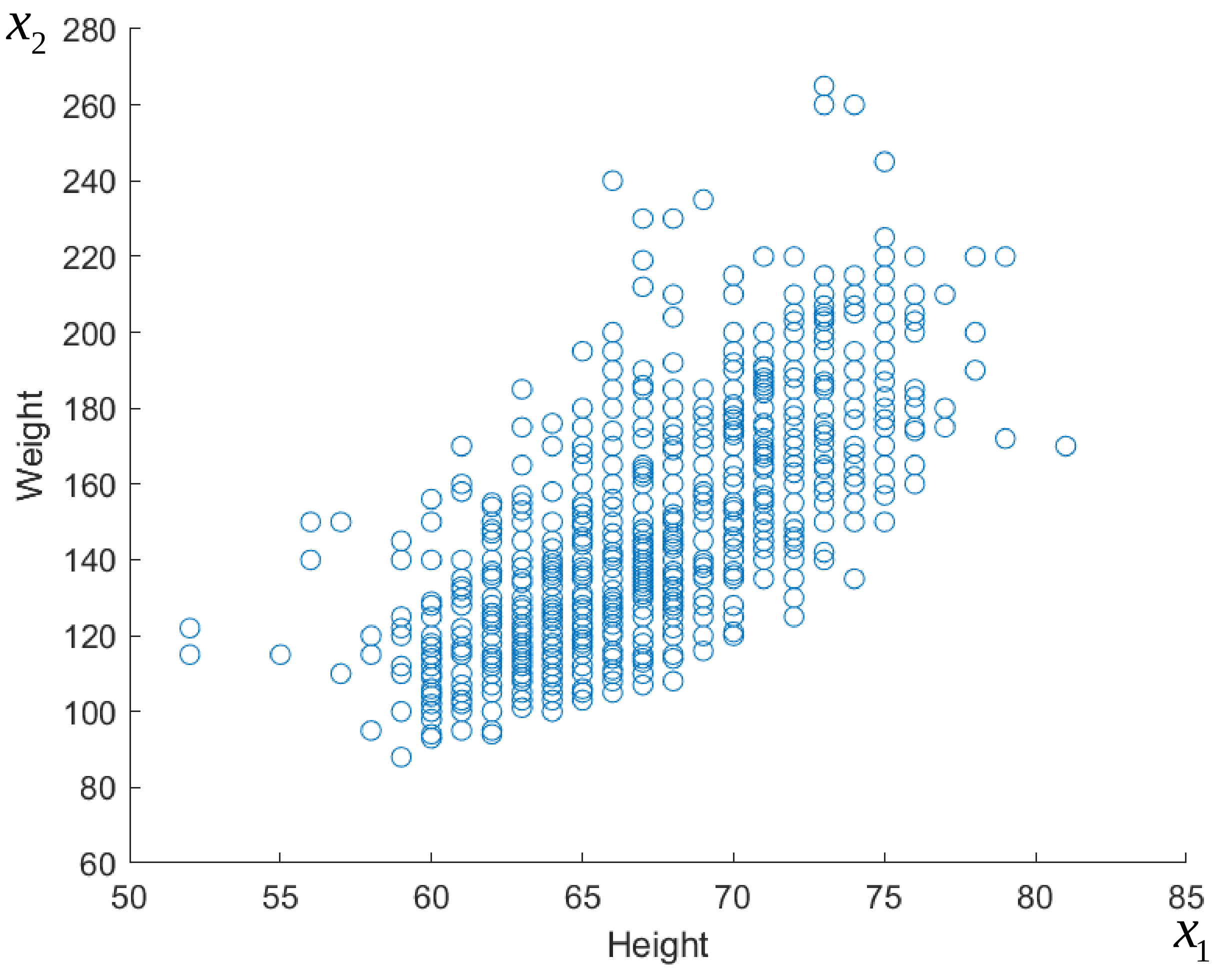
Dr. Rowe kindly asked his previous introductory statistics classes to voluntarily provide their

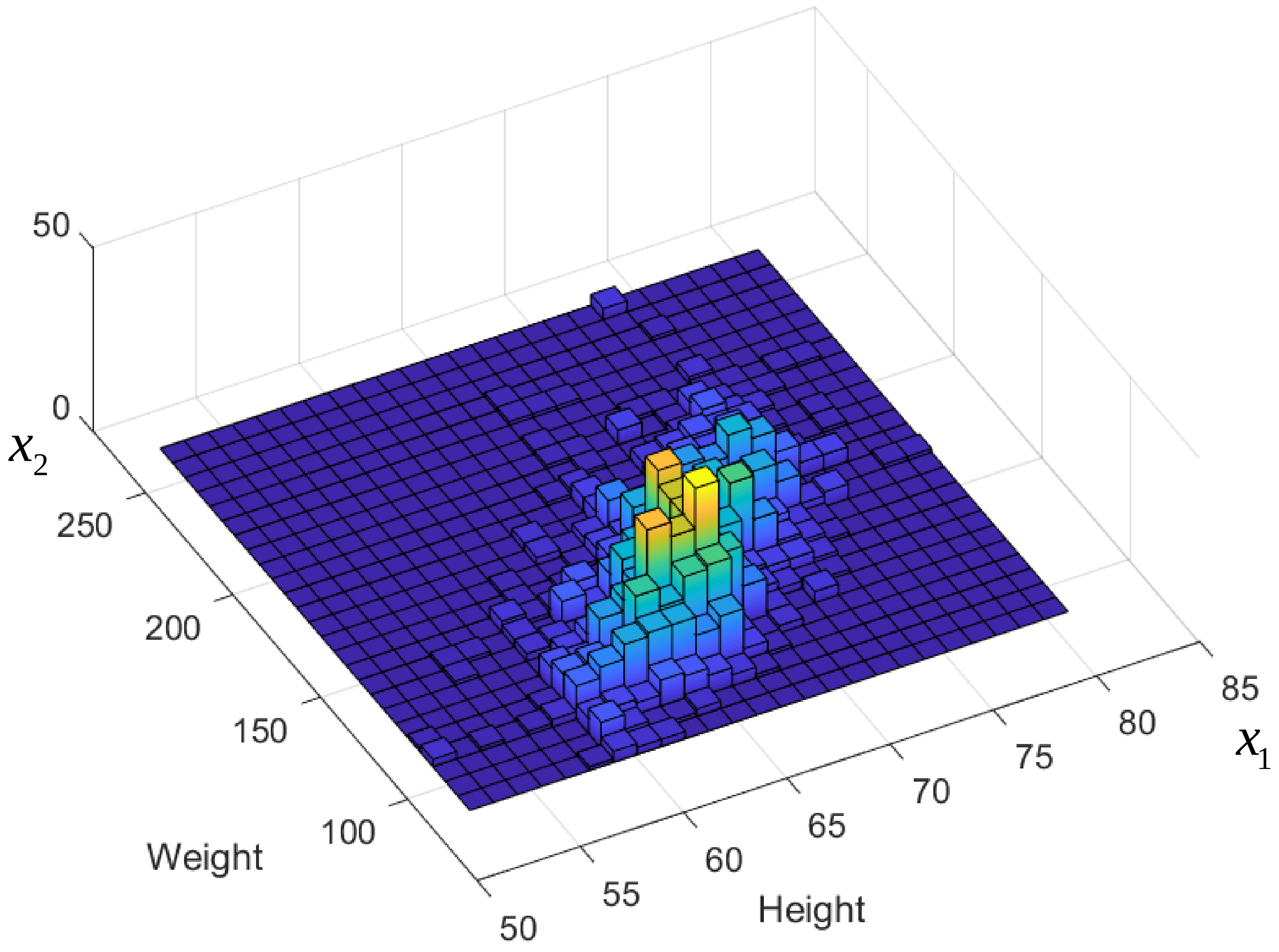
x_1=height
x_2=weight
x_3=birth gender.

This was done for 16 classes.

There were n=1366 usable data points.

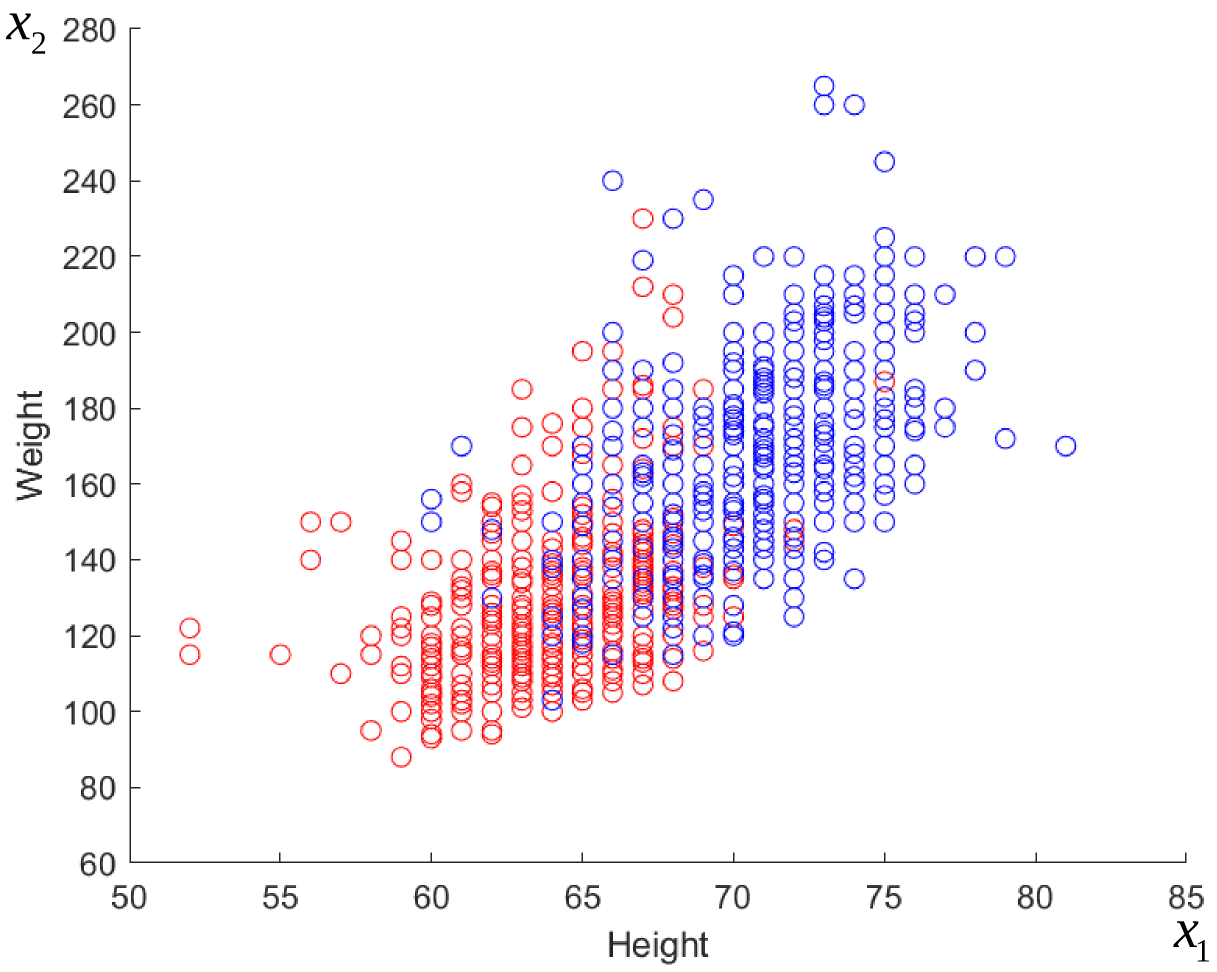
Motivation

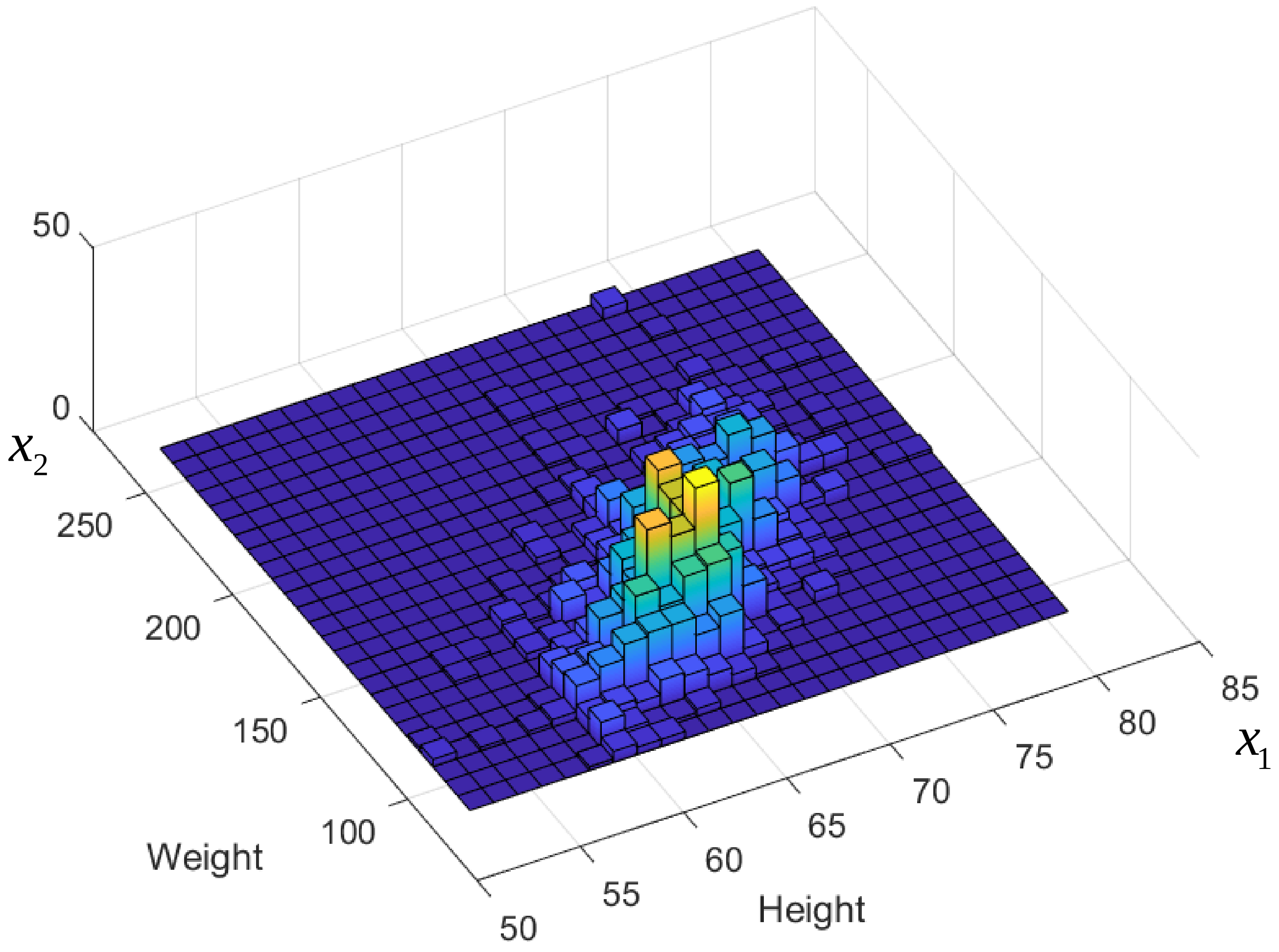




$$\underset{2\times 1}{x} = \begin{pmatrix} height \\ weight \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Motivation





○ male

○ female

$$\underset{2\times 1}{x} = \begin{pmatrix} height \\ weight \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Properties and Summaries

Assume that the 2D continuous random variable (RV) $x=(x_1,x_2)'$ can take on values

$x_1 \in (a_1,b_1) \quad x_2 \in (a_2,b_2)$

then, the probability density function (PDF) is given by

$f(x|\theta)$ defined for $x_1 \in (a_1,b_1) \quad x_2 \in (a_2,b_2)$

where x can be defined within a 2D infinite interval

and θ are any parameters that the PDF depends on.

Properties and Summaries

A bivariate (2D) PDF

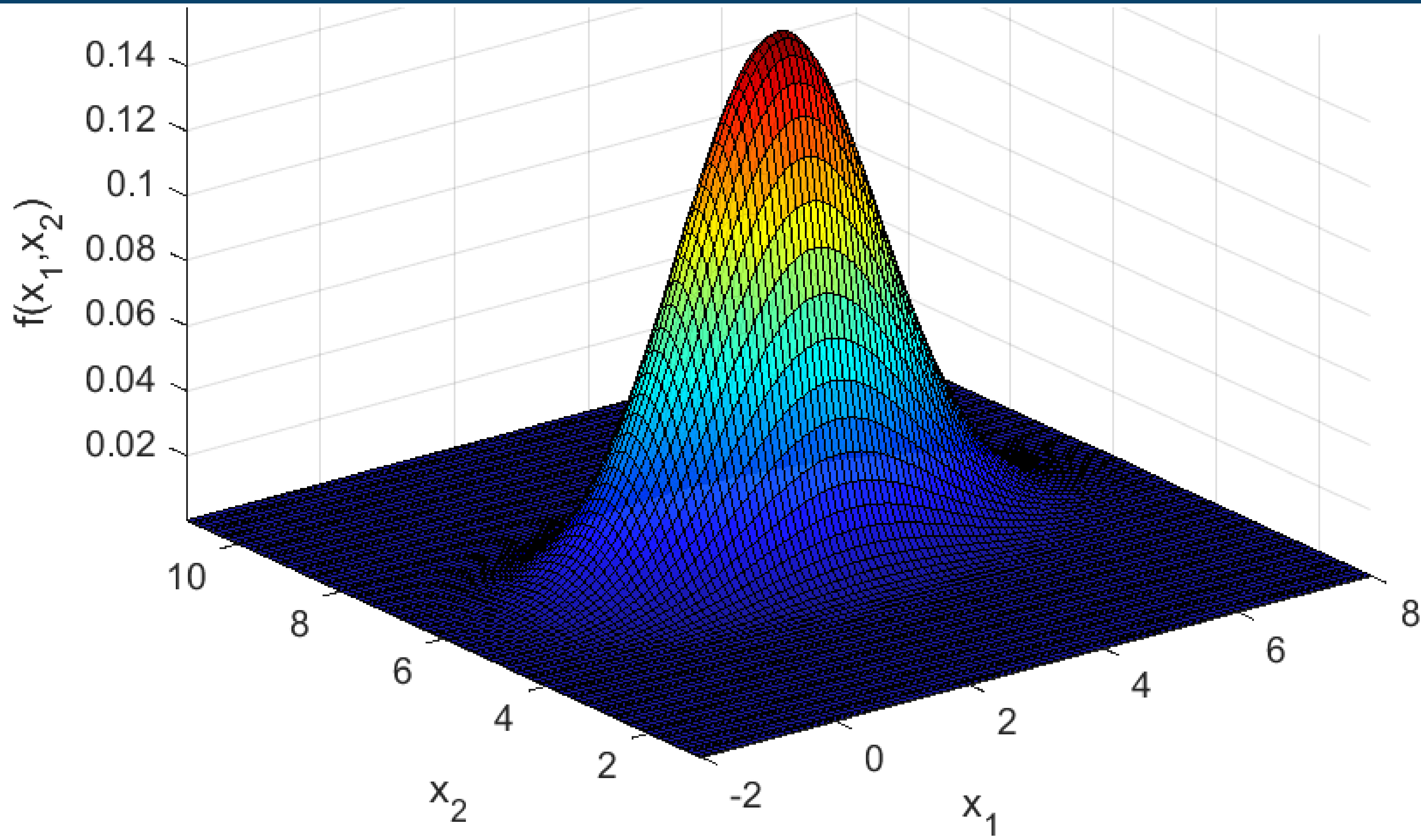
of two continuous random

variables (x_1,x_2) depending

upon parameters θ satisfies

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

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



Given θ, we completely know $f(x_1,x_2|\theta)$.

Properties and Summaries

Let $\underset{2\times 1}{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be a 2-dimensional (or p-dimensional)

random variable with PDF of $\underset{2\times 1}{x}$ being $f(x|\theta)$, then

$$E(x|\theta) = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \underset{2\times 1}{\mu} \qquad \operatorname{cov}(x|\theta) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} = \underset{2\times 2}{\Sigma}$$

which come from marginal PDFs.

Properties and Summaries

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

The marginal PDFs of x_1 and x_2 are

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

$$f(x_1|\theta) = \int_{x_2} f(x_1, x_2|\theta) dx_2$$

$$f(x_2|\theta) = \int_{x_1} f(x_1, x_2|\theta) dx_1$$

With marginal means

$$\mu_1 = \int_{x_1} x_1 f(x_1|\theta) dx_1$$

$$\mu_2 = \int_{x_2} x_2 f(x_2|\theta) dx_2$$

Properties and Summaries

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

And with marginal variances and covariance

$$\mathrm{cov}(x|\theta) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

$$\sigma_1^2 = \int_{x_1} (x_1 - \mu_1)^2 \int_{x_2} f(x_1, x_2 | \theta) dx_2 dx_1 = \int_{x_1} (x_1 - \mu_1)^2 f(x_1 | \theta) dx_1$$

$$\sigma_2^2 = \int_{x_2} (x_2 - \mu_2)^2 \int_{x_1} f(x_1, x_2 | \theta) dx_1 dx_2 = \int_{x_2} (x_2 - \mu_2)^2 f(x_2 | \theta) dx_2$$

$$\sigma_{12} = \int_{x_1} \int_{x_2} (x_1 - \mu_1)(x_2 - \mu_2) f(x_1, x_2 | \theta) dx_1 dx_2$$

Properties and Summaries

Conditional PDFs can also be found

$$f(x_1 \mid x_2, \theta) = \frac{f(x_1, x_2 \mid \theta)}{f(x_2 \mid \theta)}$$

conditional PDFs

marginal PDF

joint PDFs

$$f(x_2 \mid x_1, \theta) = \frac{f(x_1, x_2 \mid \theta)}{f(x_1 \mid \theta)}$$

marginal PDF

And since they are PDFs, conditional summary measures can also be found.

Bivariate Normal PDFs

The bivariate normal PDF can be written as

$$f_X(x_1, x_2 \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}Q\right]$$

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

$$\sigma_1, \sigma_2 > 0, \ -1 < \rho < 1 \qquad -\infty < x_1, x_2, \mu_1, \mu_2 < \infty$$

$$\rho = \sigma_{12} / (\sigma_1\sigma_2) \qquad \sigma_{12} = \mathrm{cov}(x_1, x_2)$$

Bivariate Normal PDFs

$$\sigma_1 > 0,\ \ \sigma_2 > 0$$

If there is no correlation between x_1 and x_2, ρ=0 then

$$f_X(x_1, x_2 \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{ -\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1} \right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2} \right)^2 \right] \right\}$$

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

$$-\infty < x_1, x_2, \mu_1, \mu_2 < \infty$$

$$\sigma_1, \sigma_2 > 0$$

Bivariate Normal PDFs

$$\sigma_1 > 0,\ \ \sigma_2 > 0$$

If there is no correlation between x_1 and x_2, ρ=0 then

$$f_X(x_1, x_2 \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{ -\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 \right] \right\}$$

$$f_X(x_1, x_2 \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{(x_1-\mu_1)^2}{2\sigma_1^2} \right] \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left[-\frac{(x_2-\mu_2)^2}{2\sigma_2^2} \right]$$

$$f_X(x_1, x_2 \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = f_{X_1}(x_1 \mid \mu_1, \sigma_1^2)\, f_{X_2}(x_2 \mid \mu_2, \sigma_2^2)$$

Bivariate Normal PDFs

It can be shown that

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

$$Q = \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}' \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}$$

$$Q = (x-\mu)'\Sigma^{-1}(x-\mu)$$

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

Bivariate Normal PDFs

It can be shown that

$$|\Sigma| = \begin{vmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_1^2 \end{vmatrix}$$

$$= \sigma_1^2 \sigma_2^2 - \sigma_{12}\sigma_{12}$$

$$= \sigma_1^2 \sigma_2^2 - \rho^2 \sigma_1^2 \sigma_2^2 \qquad \rho = \sigma_{12} / (\sigma_1 \sigma_2)$$

$$= \sigma_1^2 \sigma_2^2 (1 - \rho^2)$$

Bivariate Normal PDFs

$$|\Sigma| = \sigma_1^2\sigma_2^2(1-\rho^2)$$

$$Q = (x-\mu)'\Sigma^{-1}(x-\mu)$$

This means that

$$f_X(x_1, x_2 \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}Q\right]$$

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

becomes

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

$$x, \mu \in \mathbb{R}^2$$

$$\Sigma > 0$$

$\uparrow$ set of pos def matrices

$$-1 < \rho < 1$$

which is the parameterization that Statisticians use.

Bivariate Normal PDFs

If a random variable $\underset{2\times 1}{x}$ has a normal distribution with

mean vector $\underset{2\times 1}{\mu}$ and variance-covariance matrix $\underset{2\times 2}{\Sigma}$, then

mean vector mean vector

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

covariance matrix covariance matrix

$x,\mu\in\mathbb{R}^2$

$p=2$

$\Sigma>0$

set of pos def matrices

mean vector covariance matrix

and we write $\underset{2\times 1}{x}\sim N(\underset{2\times 1}{\mu},\underset{2\times 2}{\Sigma})$. The covariance matrix $\underset{2\times 2}{\Sigma}$, has to

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

make sure you know what this means

Bivariate Normal PDFs

$$|\Sigma| = \sigma_1^2\sigma_1^2(1-\rho^2)$$

From the bivariate normal PDF,

$$Q = (x-\mu)'\Sigma^{-1}(x-\mu)$$

$$f(\underset{2\times1}{x}|\underset{2\times1}{\mu},\underset{2\times2}{\Sigma}) = (2\pi)^{-2/2}|\Sigma|^{-1/2}\, e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}$$

The marginal normal PDFs of x_1 and x_2 can be shown to be

$$f(x_i|\mu_i,\sigma_i^2) = \frac{e^{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}}}{\sqrt{2\pi\sigma_i^2}}$$

$$-\infty < x_i < \infty$$
$$-\infty < \mu_i < \infty$$
$$\sigma_i > 0$$

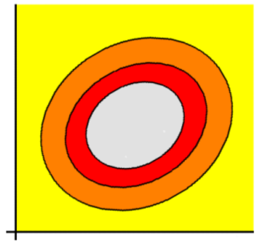
$$i = 1,2$$

with marginal mean μ_i and variance σ_i.

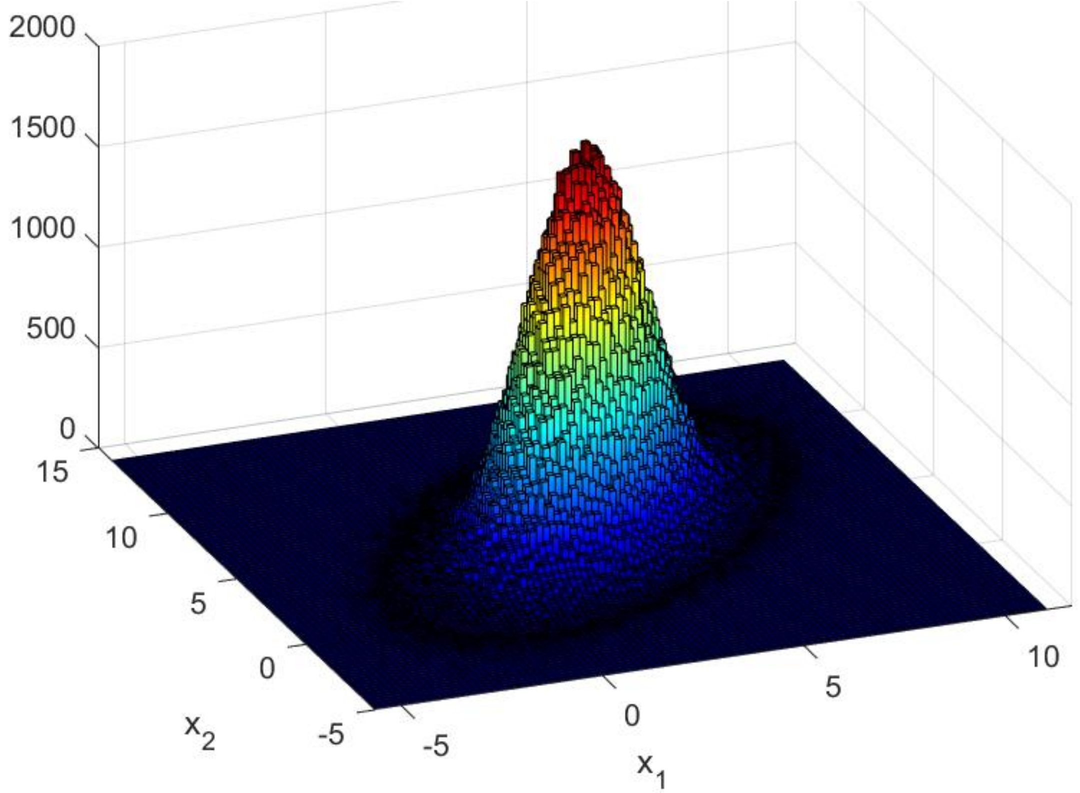
Bivariate Normal PDFs

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

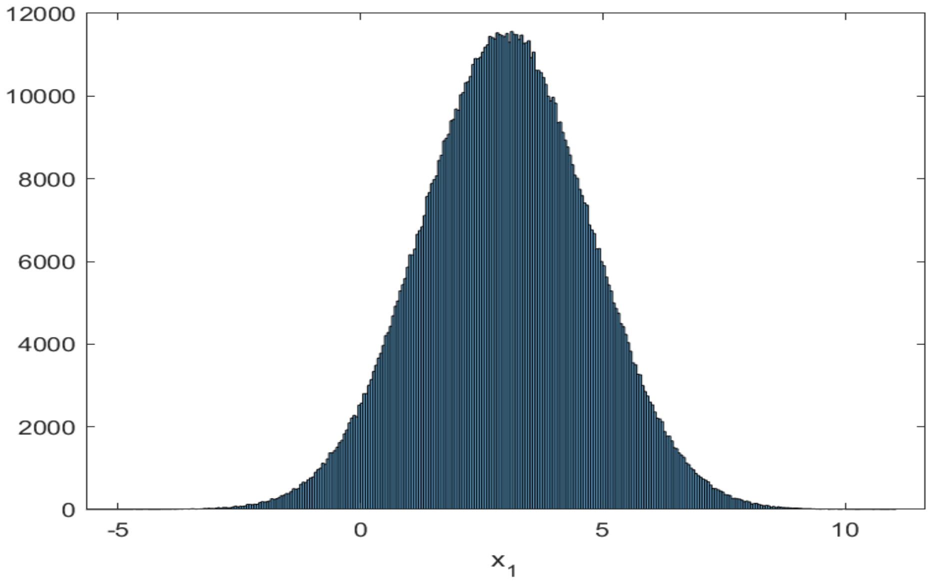
$$\underset{2\times 1}{\mu}=\begin{pmatrix}3\\5\end{pmatrix}\qquad \underset{2\times 2}{\Sigma}=\begin{pmatrix}3 & 2\\2 & 4\end{pmatrix}$$

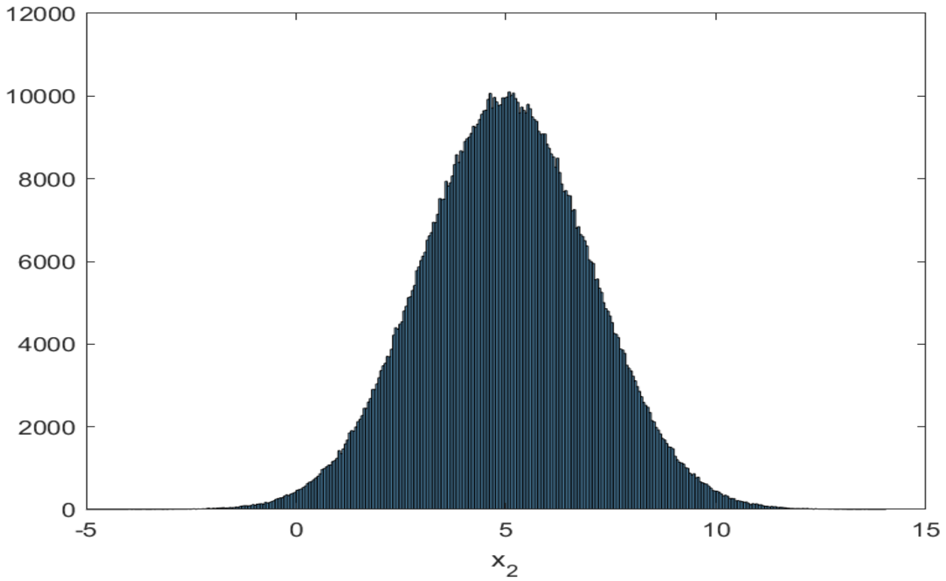


cross section



marginals are normal





Bivariate Normal PDFs

```
rng('default')
n=10^6;
mu=[3,5];,Sigma=[3,2;2,4];
X = mvnrnd(mu,Sigma,n);

figure;
hist3(X,[100,100],'CDataMode','auto','FaceColor','interp','EdgeColor',[0,0,0])
xlim([mu(1,1)-5*sqrt(3) mu(1,1)+5*sqrt(3)]),ylim([mu(1,2)-5*sqrt(4) mu(1,2)+5*sqrt(4)])
xlabel('x_1'), ylabel('x_2')
colormap(jet)

figure;
histogram(X(:,1))
xlim([mu(1,1)-5*sqrt(3) mu(1,1)+5*sqrt(3)])
xlabel('x_1')

figure;
histogram(X(:,2))
xlim([mu(1,2)-5*sqrt(4) mu(1,2)+5*sqrt(4)])
xlabel('x_2')
```

Bivariate Normal PDFs

$$|\Sigma| = \sigma_1^2\sigma_1^2(1-\rho^2)$$

And from the bivariate normal PDF,

$$Q = (x-\mu)'\Sigma^{-1}(x-\mu)$$

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

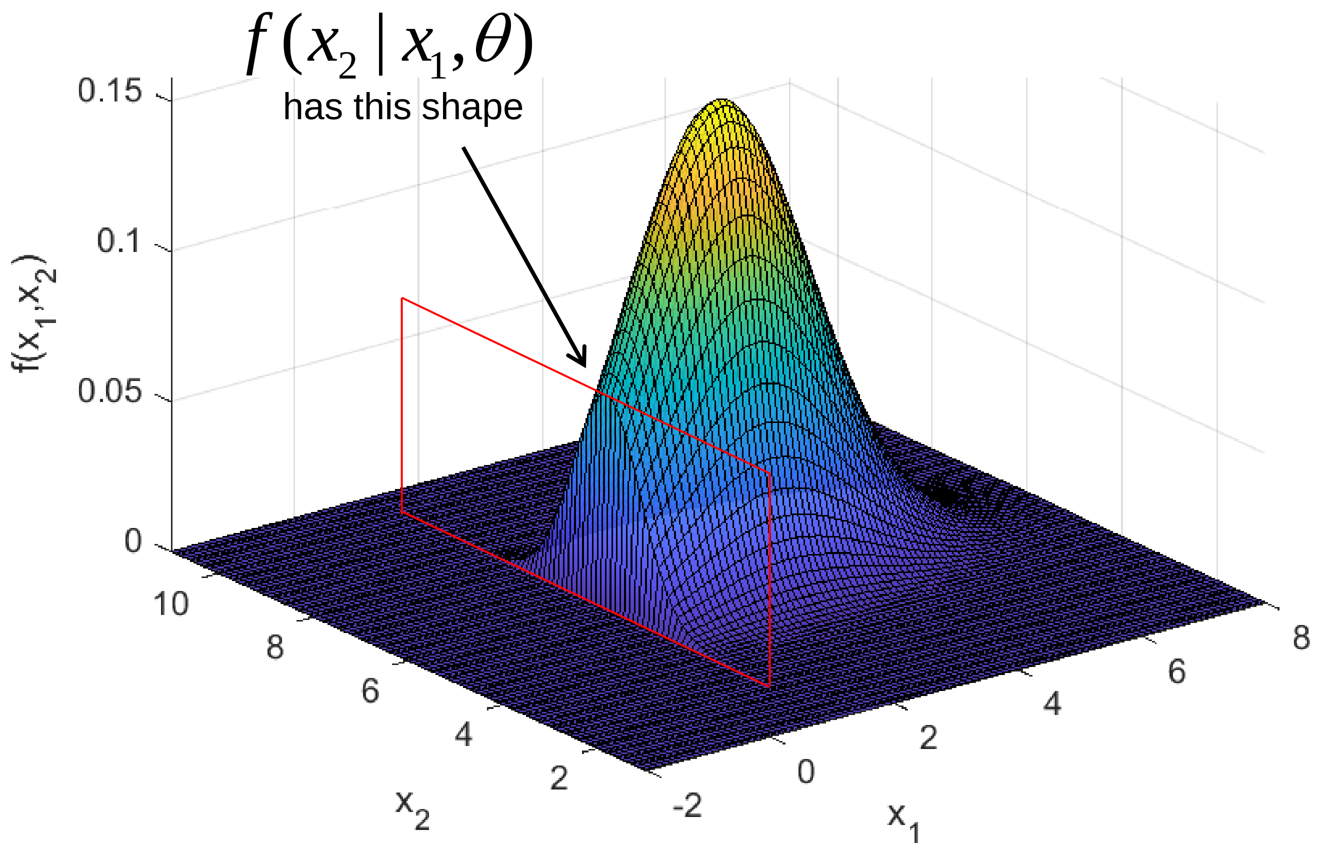
The conditional PDFs of $x_2|x_1$ can be shown to be

$$f(x_2|x_1,\mu_1,\sigma_1^2,\mu_2,\sigma_2^2,\rho) = \frac{\exp\left\{-\frac{\left[x_2-\mu_2-\frac{\rho\sigma_2}{\sigma_1}(x_1-\mu_1)\right]^2}{2\sigma_2^2(1-\rho^2)}\right\}}{(1-\rho^2)^{1/2}\sqrt{2\pi\sigma_2^2}}\,, \qquad \begin{array}{c} -\infty < x_i < \infty \\ -\infty < \mu_i < \infty \\ \sigma_i > 0 \end{array}\,.$$

Note that becomes marginal when $\rho=0$.
Can also compute conditional PDF summaries.

Bivariate Normal PDFs

The conditional PDF of $x_2|x_1$ has the general slice shape but normalized to integrate to 1.



$$f(x_2 \mid x_1, \mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho) = \frac{\exp\left\{ -\frac{\left[x_2 - \mu_2 - \frac{\rho\sigma_2}{\sigma_1}(x_1 - \mu_1) \right]^2}{2\sigma_2^2(1-\rho^2)} \right\}}{(1-\rho^2)^{1/2}\sqrt{2\pi\sigma_2^2}}$$

$$-\infty < x_i < \infty$$
$$-\infty < \mu_i < \infty$$
$$\sigma_i > 0$$

Bivariate Normal PDFs

Theorem:

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

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

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

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

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

Bivariate Normal PDFs

Theorem: Marginal

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

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

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

and partition $\underset{p\times 1}{x} = \begin{pmatrix} x_A \\ x_B \end{pmatrix} \begin{matrix} \}\ p_A\times 1 \\ \}\ p_B\times 1 \end{matrix}$, $\underset{p\times 1}{\mu} = \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix} \begin{matrix} \}\ p_A\times 1 \\ \}\ p_B\times 1 \end{matrix}$, $\underset{p\times p}{\Sigma} = \begin{pmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \underbrace{\Sigma_{BA}}_{p_A\times 1} & \underbrace{\Sigma_{BB}}_{p_B\times 1} \end{pmatrix} \begin{matrix} \}\ p_A\times 1 \\ \}\ p_B\times 1 \end{matrix}$

where $p_A+p_B=p$, then the marginal PDFs of $\underset{p_A\times 1}{x_A}$ and $\underset{p_B\times 1}{x_B}$ are

$x_A \sim N(\mu_A,\Sigma_{AA})$ and $x_B \sim N(\mu_B,\Sigma_{BB})$!

Bivariate Normal PDFs

Theorem: Conditional

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

$$E(x|\mu,\Sigma) = \underset{p\times 1}{\mu} \quad \text{and} \quad cov(x|\mu,\Sigma) = \underset{p\times p}{\Sigma}$$

think of p=2

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

and partition $\underset{p\times 1}{x} = \begin{pmatrix} x_A \\ x_B \end{pmatrix} \begin{matrix} \}\ p_A\times 1 \\ \}\ p_B\times 1 \end{matrix}$, $\underset{p\times 1}{\mu} = \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix} \begin{matrix} \}\ p_A\times 1 \\ \}\ p_B\times 1 \end{matrix}$, $\underset{p\times p}{\Sigma} = \begin{pmatrix} \underbrace{\begin{matrix}\Sigma_{AA} \\ \Sigma_{BA}\end{matrix}}_{p_A\times 1} & \underbrace{\begin{matrix}\Sigma_{AB} \\ \Sigma_{BB}\end{matrix}}_{p_B\times 1} \end{pmatrix} \begin{matrix} \}\ p_A\times 1 \\ \}\ p_B\times 1 \end{matrix}$

where $p_A+p_B=p$, then the conditional PDFs of $\underset{p_A\times 1}{x_A}$ and $\underset{p_B\times 1}{x_B}$ are

$x_A|x_B \sim N(\mu_A+\Sigma_{AB}\Sigma_{BB}^{-1}(x_B-\mu_B), \Lambda_{AA})$ and $x_B|x_A \sim N(\mu_B+\Sigma_{BA}\Sigma_{AA}^{-1}(x_A-\mu_A), \Lambda_{BB})$

where $\Lambda_{BB}= \Sigma_{BB}-\Sigma_{BA}\Sigma_{AA}^{-1}\Sigma_{AB}$ and $\Lambda_{AA}= \Sigma_{AA}-\Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$.

Discussion

Questions?

Homework 4

1. Assume Marquette Undergrads heights have true μ_h=67 in and σ_h=2 in while their true weights have μ_w=150 lbs and σ_w=4 lbs with ρ=.75.

a) Generate 10^4 h/w 2×1 vectors from a bivariate normal PDF.

b) Calculate sample means, variances, and covariance. Compare to truth.

c) Make an (x_1, x_2) bivariate histogram and marginal histograms for x_1 and x_2. (For marginal just ignore other.)

d) Plot the theoretical conditional distribution of weights x_2 given a height of x_1=69 in.

e*) For an interval of heights $69\ in \pm 0.5\ in$ make a histogram. Compare.

* For students in 5790.

Homework 4

2*. Derive by pencil and paper the theoretical marginal PDF

$f(x_1|\mu_1,\sigma_1,\mu_2,\sigma_2,\rho)$ of heights x_1 given all the parameters.

Start with equation for bivariate.

3*. Derive by pencil and paper the theoretical conditional PDF

$f(x_2|x_{1,}\mu_1,\sigma_1,\mu_2,\sigma_2,\rho)$ of weights x_2 given a height of x_1 and all the parameters.

Start with equation for bivariate.

4* Superimpose plots of theoretical PDFs on histograms in #1.

* For students in 5790.