ASEN 5070
Statistical Orbit determination I
Fall 2012

Professor George H. Born
Professor Jeffrey S. Parker

Lecture 7: Spaceflight Ops and Statistics

Announcements

- Homework 1 Graded
  - Comments included on D2L
  - Any questions, talk with us this week
- Homework 2 CAETE due Thursday
  - Graded soon after
- Homework 3 due Thursday
- Homework 4 out today

Lecture Plan

- Review Homework and Quizzes
- Finish Mars Odyssey
- Review Statistics
- Start on some serious Stat OD!

Homework 3

- Questions yet?

Homework 4

1. Two random variables have the joint density function given by:

$$
\begin{aligned}
f(x, y) &= k(x^2 + y^2), && 0 \le x \le 2, \quad 1 \le y \le 3 \\
f(x, y) &= 0, && \text{elsewhere}
\end{aligned}
$$

a) Find k

b) Find $p(1 < x \le 2,\ 2 < y \le 3)$

c) Find $p(1 \le x \le 2)$

d) Find $p(x + y \ge 4)$

e) Find $p(x + y = 4)$

f) Find $p(x \le 1 \ / \ y = 3)$

g) Find σ_x

h) Find $p(1 < x < 2 \ / \ 1 < y < 2)$

Homework 4

2. Show that the moment generating function for the univariate normal distribution

$$f(x) = \frac{1}{b\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left(\frac{x-a}{b}\right)^2\right]} \qquad -\infty \leq x \leq \infty$$

is given by

$$M_x(\theta) = e^{\left[\frac{\theta^2 b^2}{2} + a\theta\right]}$$

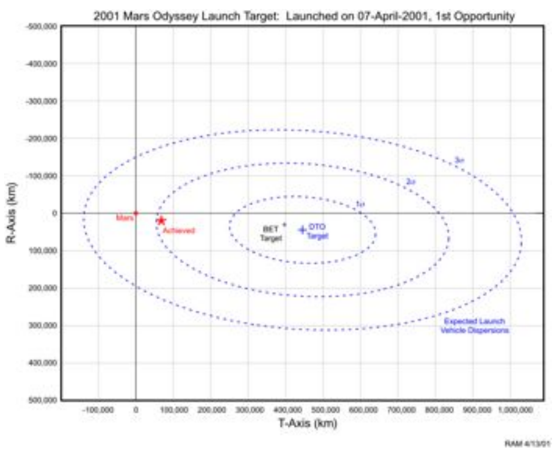
3. If x and y are independent random variables, show that

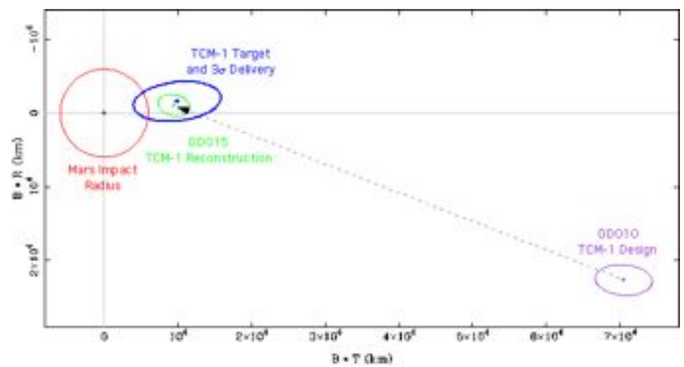
$$\sigma^2(xy) = \sigma^2(x)\sigma^2(y) + \lambda^2(x)\sigma^2(y) + \lambda^2(y)\sigma^2(x)$$

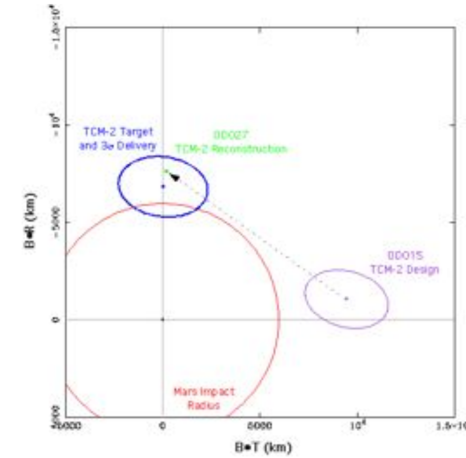
using the notation of Appendix A:

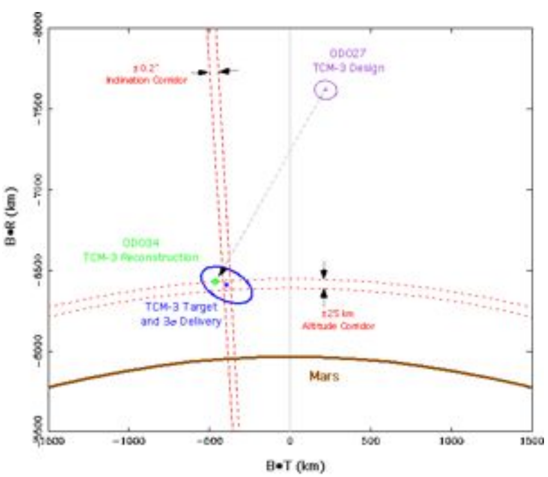
$$\sigma^2(xy) = E\left[xy - E(xy)\right]^2, \ \sigma^2(x) = \mu_{20}, \ \sigma^2(y) = \mu_{02}, \ \lambda(x) = \lambda_{10}, \ \lambda(y) = \lambda_{01}$$

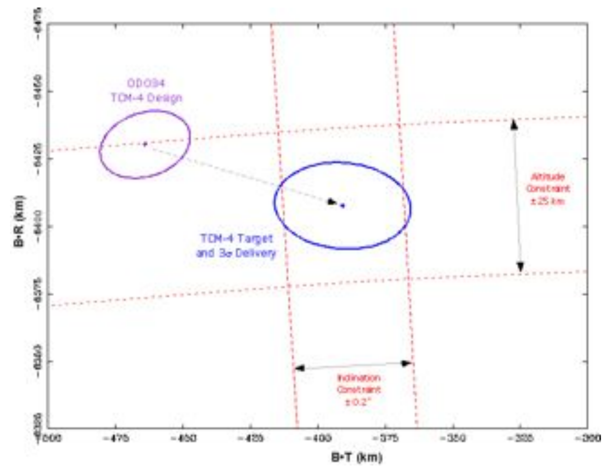
Mars Odyssey

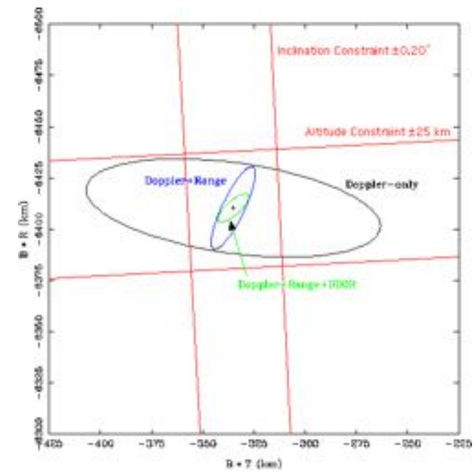




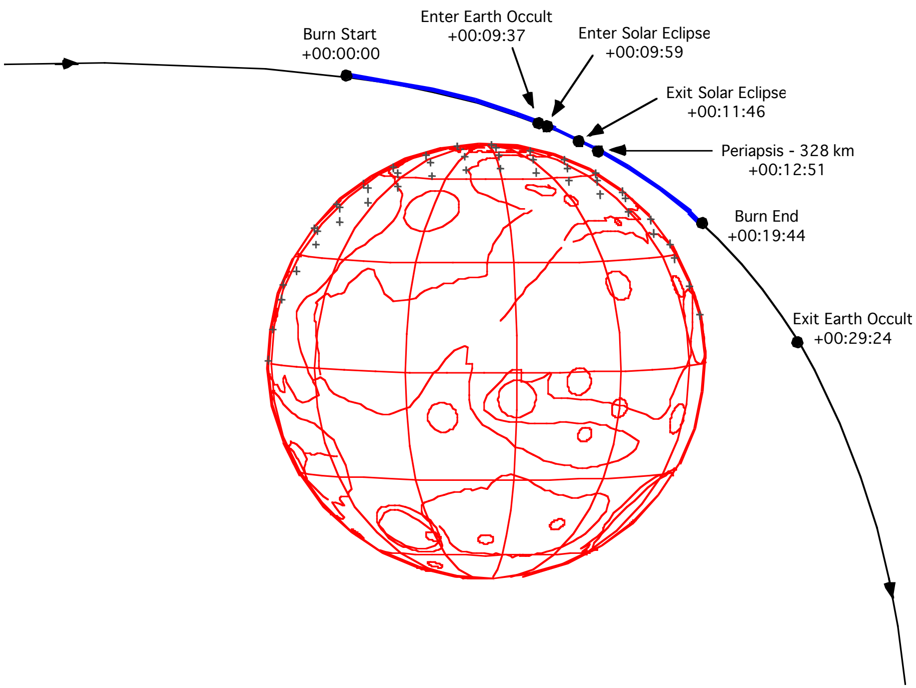


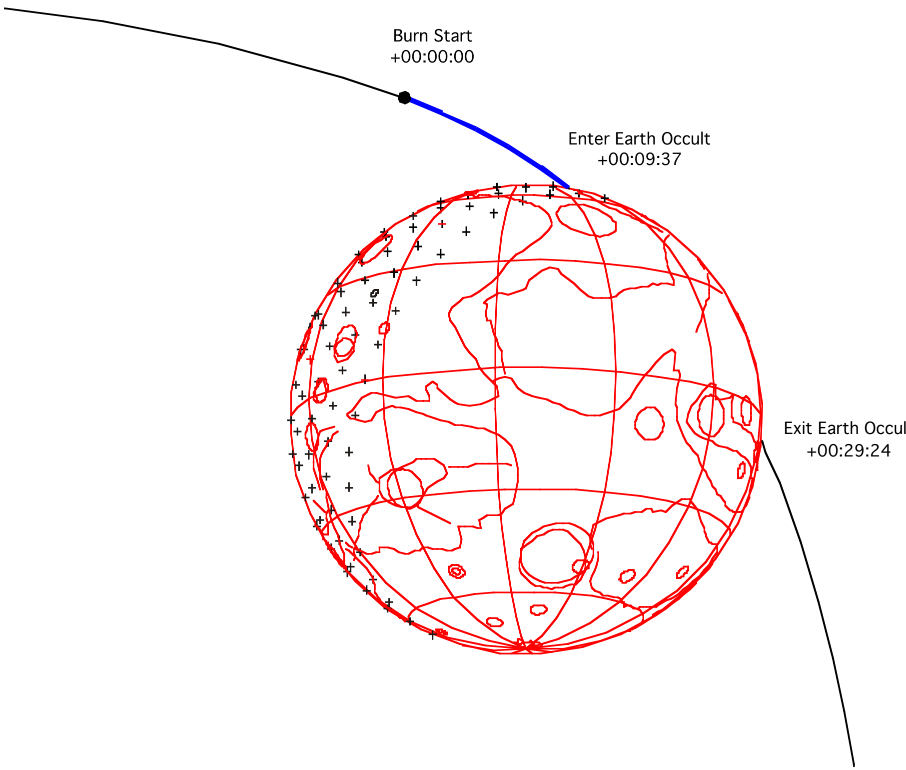






Mars Odyssey





Aerobraking Nav Prediction Accuracy

- Requirement
  - Must predict Periapsis Time to within 225 sec
  - Must predict Periapsis Altitude to within 1.5 km

- Capability
  - Altitude requirement easily met with MGS gravity field (Nav Plan)
  - Timing requirement uncertainty dominated by assumption on future drag pass atmospheric uncertainty

- Atmospheric Variability
  - Total Orbit-to-Orbit Atmospheric variability: 80% (MGS: 90%)

- Periapsis timing prediction
  - To first order, the expected change in orbit period per drag pass will indicate how well future periapses can be predicted
  - This simplifying assumption is supported by OD covariance analysis

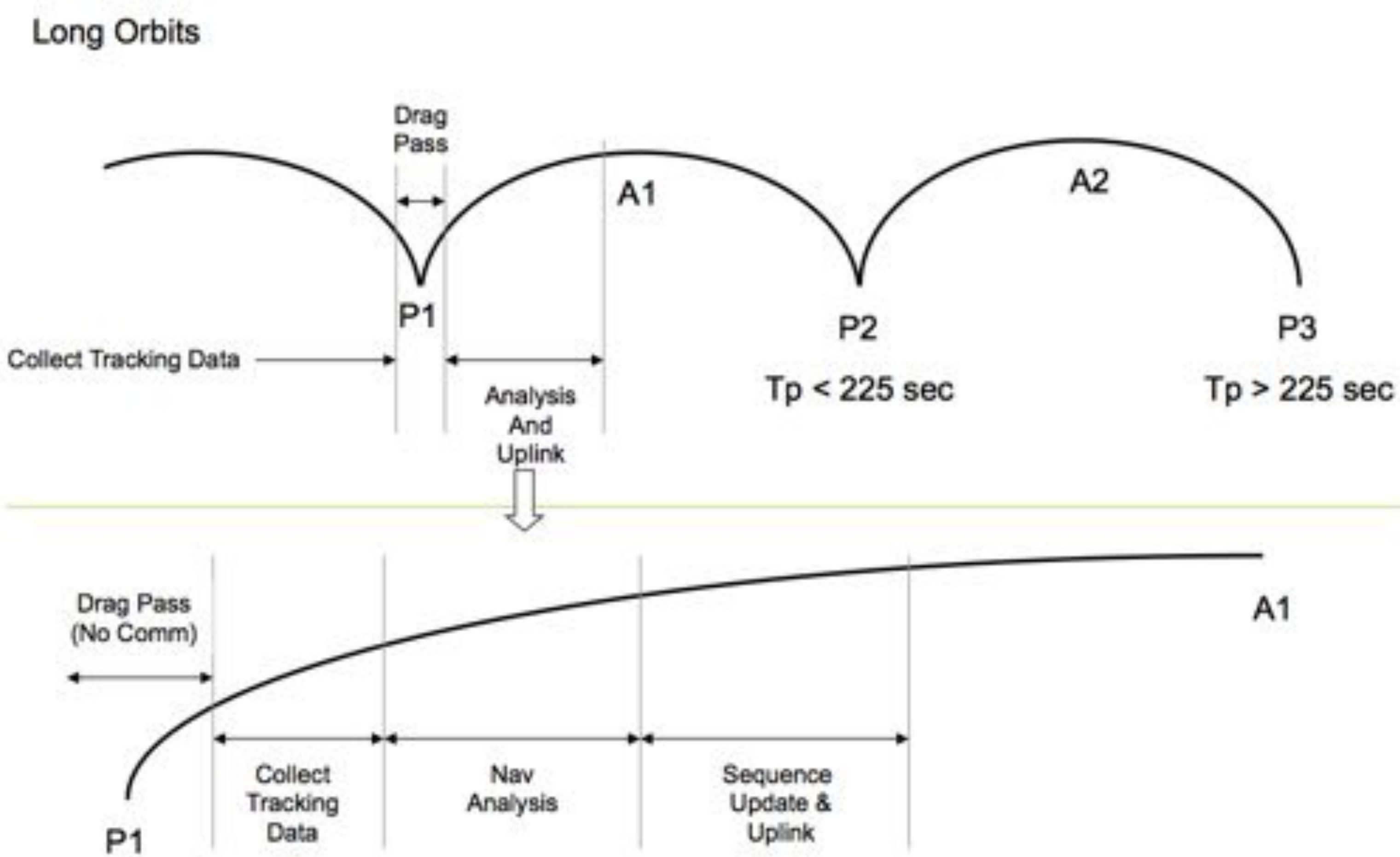
Nav Predict Capability

- Example
  - Total expected Period change for a given drag pass is 1000 seconds
  - Atmosphere could change density by 80%
  - Resulting Period change could be off by 80% = 800 sec
  - If orbit Period is different by 800 seconds, then the time of the next periapsis will be different by 800 seconds
  - This fails to meet the 225 sec requirement

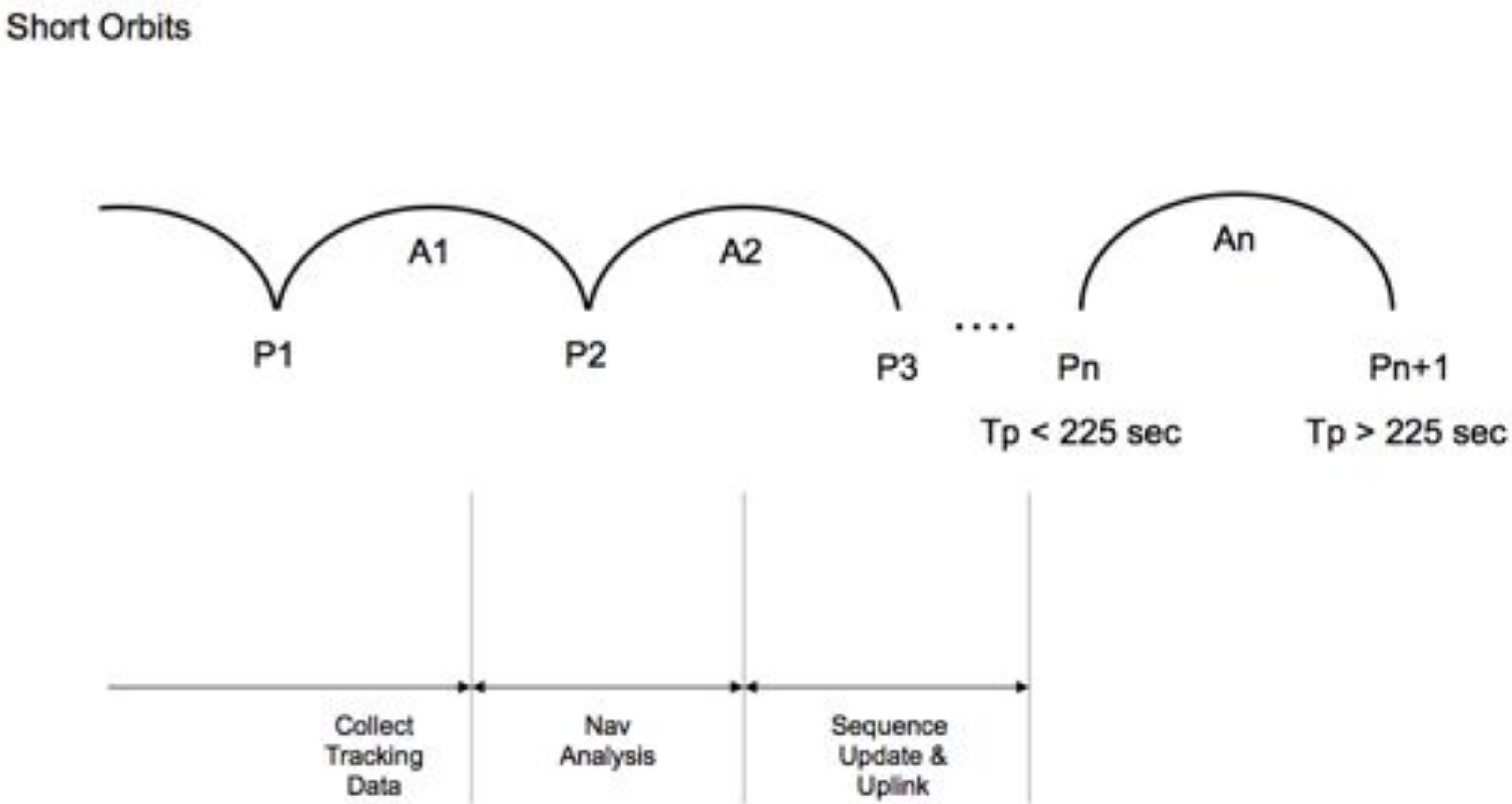
- Large Period Orbits
  - Period change per rev is large
  - Therefore can never predict more than 1 periapsis ahead within the 225 sec requirement with any confidence

- Small Period Orbits
  - Period change per rev is small (for example 30 seconds)
  - Therefore can predict several periapses in the future to within the 225 second requirement
  - Example: 80% uncertainty (24 sec) will allow a 9 rev predict

Aerobraking Navigation Process



Aerobraking Navigation Process



What contributed to ODY success?

- A Baseline set of Navigation solution strategies were identified
  - Varied data arcs, data types, data weights, parameter estimates, a-prioris
- These solutions were regularly performed and trended
  - Built a time history of trajectory solutions
  - Trended evolution of parameter estimates and encounter conditions
  - Lessons learned from MCO and MPL
- Regularly demonstrate consistency to Project and NAG
  - Weekly Status Reports
  - Daily Status after TCM-4 (MOI-12 days) "Daily Show"
- Shadow navigators
  - Independent solutions run by Sec312 personnel (Bhaskaran, Portock)

Mars Odyssey

- Questions
- Quick Break
- Next up:
  - Statistics
  - Stat OD

The Variance-Covariance Matrix

The correlation coefficient is defined as

$$\rho_{XY} \equiv \frac{E\{[X - E(X)]\,[Y - E(Y)]\}}{\{E[X - E(X)]^2\}^{1/2}\{E\,[Y - E(Y)]^2\}^{1/2}}$$
$$= \frac{\mu_{11}}{\sigma(X)\sigma(Y)}\,. \qquad \text{(A.14.2)}$$

The variance-covariance matrix for an n-dimensional random vector, X, can be written as

$$P = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \rho_{2n}\sigma_2\sigma_n & \cdots & \sigma_n^2 \end{bmatrix} \qquad \text{(A.14.3)}$$

where ρ_{ij} is a measure of the degree of linear correlation between X_i and X_j.

Quiz #5 Results

Question 1 (1 point)

The variance-covariance matrix (or "covariance matrix" for short) captures correlations between each of the parameters of a state.

Is the following statement true or false?: The covariance matrix is symmetric. That is, it is equal to its transpose.

- True
- False

Quiz #5 Results

Question 1 (1 point)

The variance-covariance matrix (or "covariance matrix" for short) captures correlations between each of the parameters of a state.

Is the following statement true or false?: The covariance matrix is symmetric. That is, it is equal to its transpose.

- True
- False

The Variance-Covariance Matrix

The correlation coefficient is defined as

$$
\begin{aligned}
\rho_{XY} &\equiv \frac{E\{[X - E(X)]\,[Y - E(Y)]\}}{\{E[X - E(X)]^2\}^{1/2}\{E\,[Y - E(Y)]^2\}^{1/2}} \\
&= \frac{\mu_{11}}{\sigma(X)\sigma(Y)}\,. \qquad \text{(A.14.2)}
\end{aligned}
$$

The variance-covariance matrix for an n-dimensional random vector, X, can be written as

$$
P = \begin{bmatrix}
\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\
\rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\
\vdots & & \ddots & \vdots \\
\rho_{1n}\sigma_1\sigma_n & \rho_{2n}\sigma_2\sigma_n & \cdots & \sigma_n^2
\end{bmatrix} \qquad \text{(A.14.3)}
$$

where ρ_{ij} is a measure of the degree of linear correlation between X_i and X_j.

Quiz #5 Results

Question 2 (1 point)

Is the following statement true or false?: If two parameters within a state are perfectly correlated, then the covariance matrix is invertible.

- True
- False

Quiz #5 Results

Question 2 (1 point)

Is the following statement true or false?: If two parameters within a state are perfectly correlated, then the covariance matrix is invertible.

- True
- False

The Variance-Covariance Matrix

$$P = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \rho_{2n}\sigma_2\sigma_n & \cdots & \sigma_n^2 \end{bmatrix}$$

If two parameters are perfectly correlated then $\rho_{ij} = 1$

Say we have $\rho_{12} = 1$ and all other correlations are 0.

$$P = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 & \cdots & 0 \\ \sigma_1\sigma_2 & \sigma_2^2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} \rightarrow \text{simplify to determine rank} \rightarrow \begin{bmatrix} \sigma_1 & \sigma_2 & \cdots & 0 \\ \sigma_1 & \sigma_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

Quiz #6 Results

Information

Still on the subject of statistics (since the launch distracted us last Thursday)!

Recall: f(x) = probability density function.

Question 1 (1 point)

The value of f(x) >= 0 for all x.

- True
- False

Quiz #6 Results

Information

Still on the subject of statistics (since the launch distracted us last Thursday)!

Recall: f(x) = probability density function.

Question 1 (1 point)

The value of f(x) >= 0 for all x.

- True
- False

Quiz #6 Results

Question 2 (1 point)

The integral from -infinity to +infinity of f(x)dx is equal to 1.0

- True
- False

Quiz #6 Results

Question 2 (1 point)

The integral from -infinity to +infinity of f(x)dx is equal to 1.0

- True
- False

The probability of sampling a distribution and getting *something* is equal to 100%

Quiz #6 Results

Question 3 (1 point)

The value of f(x) may be negative so long as the integral of f(x) from -infinity to +infinity of f(x)dx is equal to 1.0.

- True
- False

Quiz #6 Results

Question 3 (1 point)

The value of f(x) may be negative so long as the integral of f(x) from -infinity to +infinity of f(x)dx is equal to 1.0.

- True
- False

Quiz #6 Results

CCAR
Colorado Center for Astrodynamics Research

Question 4 (1 point)

Consider the following lists of numbers:

A: [0.0, 0.5, 1.0]   (3 evenly distributed #s from 0-1)

B: [0.0, 0.25, 0.5, 0.75, 1.0]   (5 evenly distributed #s from 0-1)

C: [0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]   (11 evenly distributed #s from 0-1)

Which of these lists has the largest variance?   (or equivalently, which has the largest standard deviation?)

- A
- B
- C
- They all have the save variance (and standard deviation)

Quiz #6 Results

CCAR
Colorado Center for Astrodynamics Research

Question 4 (1 point)

Consider the following lists of numbers:

A: [0.0, 0.5, 1.0]   (3 evenly distributed #s from 0-1)

B: [0.0, 0.25, 0.5, 0.75, 1.0]   (5 evenly distributed #s from 0-1)

C: [0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]   (11 evenly distributed #s from 0-1)

Which of these lists has the largest variance? (or equivalently, which has the largest standard deviation?)

- A
- B
- C
- They all have the save variance (and standard deviation)

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

d) $p\left(x+y\le \frac{1}{2}\right)$

e) $p\left(x+y= \frac{1}{2}\right)$

f) Whether x and y are independent

g) $p\left(0<y<\frac{1}{2} / \frac{1}{2}<x<\frac{1}{2}+dx\right)$

h) $p\left(0<y<\frac{1}{2} / x>\frac{1}{2}\right)$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

f = joint density function

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$

Example Problems in Statistics

CCAR
Colorado Center for Astrodynamics Research

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find: a) k

f = joint density function

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$

$$k\int_0^1\int_0^1 (x+2y)dydx = 1$$

$$k\left[xy+y^2\right]_0^1 = k\left[x+1\right]$$

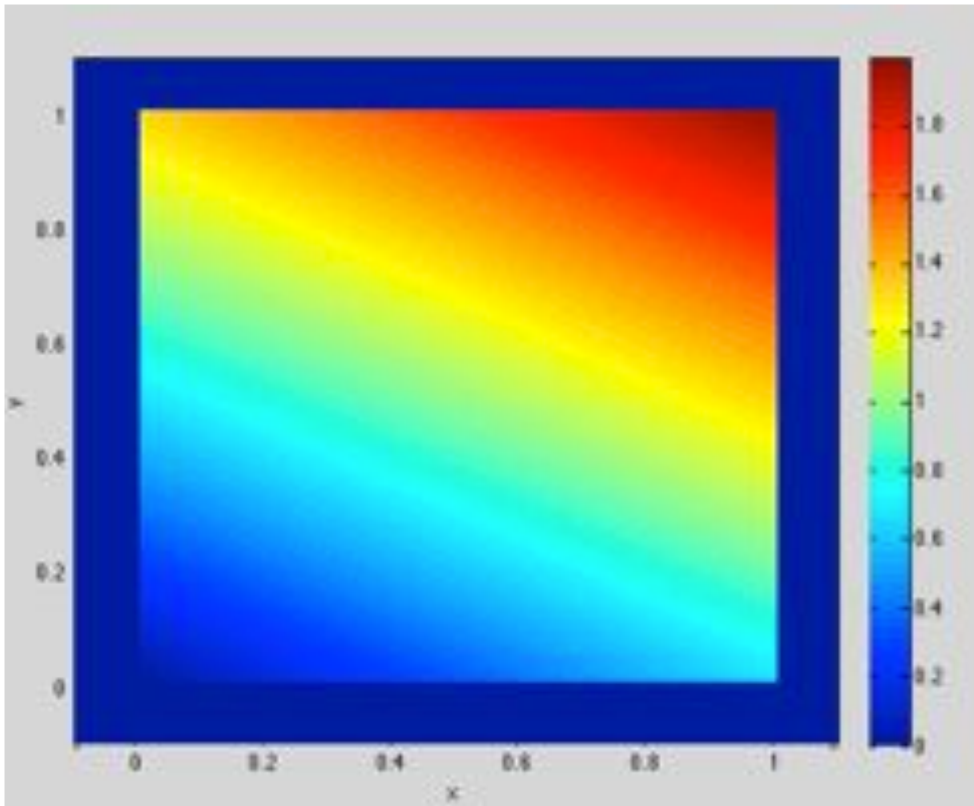
$$k\int_0^1 (x+1)dx = 1$$

$$k\left[\frac{x^2}{2}+x\right]_0^1 = k\,{}^3\!/_2 = 1 \longrightarrow k = {}^2\!/_3$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$



Example Problems in Statistics

Given

$$f(x,y) = \begin{cases} k(x+2y) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

What is the marginal density function of *x*? What does that mean?

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

What is the marginal density function of x? What does that mean?

The marginal density function of x is the probability density function of x in the presence of any y.

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & \textit{elsewhere} \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

$$g(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

We'll work this one

$$h(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

You'll work this one

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & \textit{elsewhere} \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

$$g(x)=\int_{-\infty}^{\infty} f(x,y)dy$$

$$g(x)=\frac{2}{3}\int_0^1 (x+2y)dy=\frac{2}{3}\left[xy+y^2\right]_0^1$$

$$\longrightarrow g(x)=\frac{2}{3}[x+1]$$

Example Problems in Statistics

CCAR
Colorado Center for
Astrodynamics Research

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

$$g(x)=\int_{-\infty}^{\infty} f(x,y)dy$$

$$g(x)=\frac{2}{3}\int_0^1 (x+2y)dy=\frac{2}{3}\left[xy+y^2\right]_0^1$$

$$\longrightarrow g(x)=\frac{2}{3}[x+1]$$

$$h(y)=\int_{-\infty}^{\infty} f(x,y)dx$$

???

Example Problems in Statistics

CCAR
Colorado Center for Astrodynamics Research

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of *x* and *y*

$$g(x)=\int_{-\infty}^{\infty} f(x,y)dy$$

$$h(y)=\int_{-\infty}^{\infty} f(x,y)dx$$

$$g(x)=\frac{2}{3}\int_0^1 (x+2y)dy=\frac{2}{3}\left[xy+y^2\right]_0^1$$

$$\longrightarrow g(x)=\frac{2}{3}[x+1]$$

$$h(y)=\frac{2}{3}\int_0^1 (x+2y)dx=\frac{2}{3}\left[\frac{x^2}{2}+2xy\right]_0^1$$

$$\longrightarrow h(y)=\frac{2}{3}\left[\frac{1}{2}+2y\right]$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

Example Problems in Statistics

Given

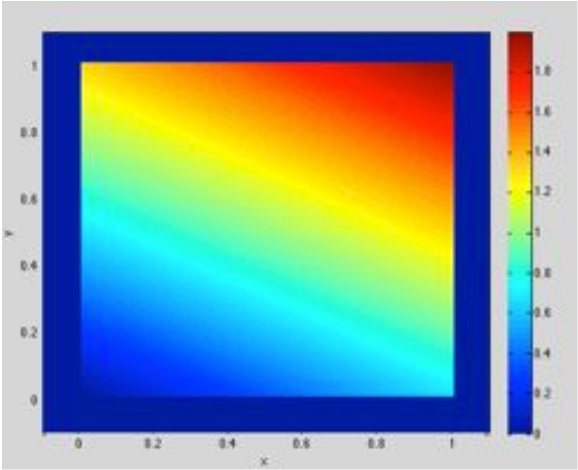
$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$



$$p\left(x\le \tfrac{1}{2}\right)=F\left(\tfrac{1}{2},\infty\right)=\int_0^{1/2} g(x)\,dx$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

$$p\left(x\le \tfrac{1}{2}\right)=F\left(\tfrac{1}{2},\infty\right)=\int_0^{1/2} g(x)\,dx$$

$$=\tfrac{2}{3}\int_0^{1/2}(x+1)\,dx \qquad =\frac{2}{3}\left[\frac{x^2}{2}+x\right]_0^{1/2} \qquad =\frac{2}{3}\left[\frac{1}{8}+\frac{1}{2}\right]=\frac{5}{12}$$

Example Problems in Statistics

CCAR Colorado Center for Astrodynamics Research

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

d) $p\left(x+y\le \frac{1}{2}\right)$

Example Problems in Statistics

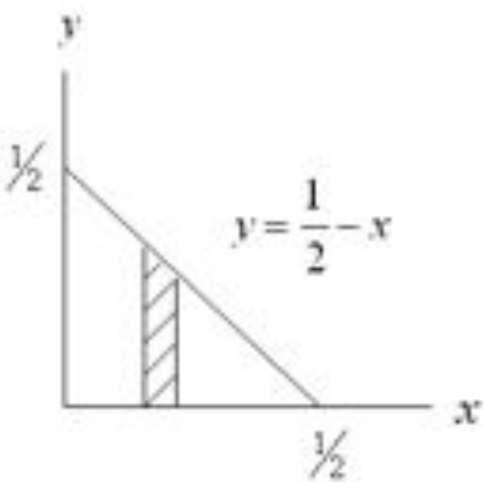
Given

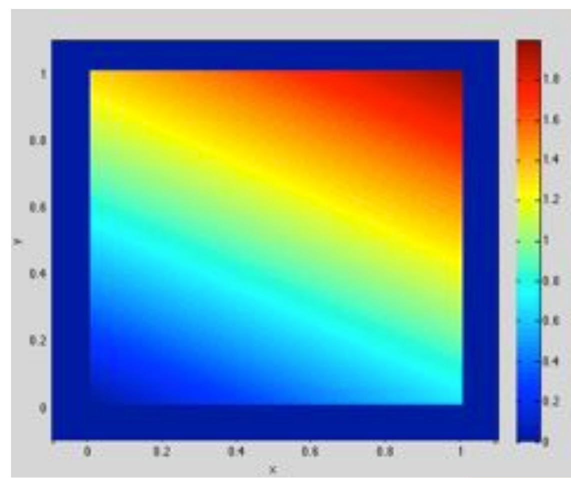
$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

d) $p\left(x+y\le \frac{1}{2}\right)$

as x ranges from $0 \to \frac{1}{2}$

and y ranges over $0 \to \frac{1}{2} - x$



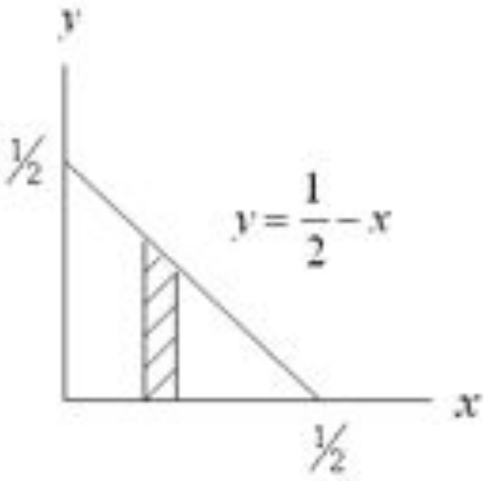


Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

d) $p\left(x+y\le \tfrac{1}{2}\right)$



as x ranges from $0\rightarrow \tfrac{1}{2}$

and y ranges over $0\rightarrow \tfrac{1}{2}-x$

$$=\frac{2}{3}\int_0^{1/2}\int_0^{1/2-x}(x+2y)\,dy\,dx$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

d) $p\left(x+y\le \tfrac{1}{2}\right)$

$$=\frac{2}{3}\int_0^{1/2}\int_0^{1/2-x}(x+2y)\,dy\,dx$$

$$=\frac{2}{3}\int_0^{1/2}\left[xy+y^2\right]_0^{1/2-x}dx$$

$$=\frac{2}{3}\int_0^{1/2}\left[-\frac{x}{2}+\frac{1}{4}\right]dx=\frac{2}{3}\left[-\frac{x^2}{4}+\frac{x}{4}\right]_0^{1/2}=\frac{1}{24}$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

d) $p\left(x+y\le \frac{1}{2}\right)$

also,

$$p\left(x+y\le \frac{1}{2}\right)=\frac{2}{3}\int_0^{1/2}\int_0^{1/2-y}(x+2y)\,dx\,dy$$

and,

$$p\left(x+y\ge \frac{1}{2}\right)=1-p\left(x+y\le \frac{1}{2}\right)$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le 1/2$

d) $p(x+y\le 1/2)$

e) $p(x+y= 1/2)$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

d) $p\left(x+y\le \frac{1}{2}\right)$

e) $p\left(x+y= \frac{1}{2}\right)$

$$p\left(x+y=\frac{1}{2}\right)=0$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

d) $p\left(x+y\le \frac{1}{2}\right)$

e) $p\left(x+y= \frac{1}{2}\right)$

f) Whether x and y are independent

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

d) $p\left(x+y\le \frac{1}{2}\right)$

e) $p\left(x+y= \frac{1}{2}\right)$

f) Whether x and y are independent

if $f(x,y) = g(x)h(y)$, then X and Y are independent

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & \textit{elsewhere} \end{cases}$$

$$g(x)=\frac{2}{3}[x+1] \qquad h(y)=\frac{2}{3}\left[\frac{1}{2}+2y\right]$$

$$f(x,y)\neq g(x)h(y), \text{ thus, } X \text{ and } Y \text{ are not independent}$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & \textit{elsewhere} \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

d) $p\left(x+y\le \frac{1}{2}\right)$

e) $p\left(x+y= \frac{1}{2}\right)$

f) Whether x and y are independent

g) $p\left(0<y<\frac{1}{2} / \frac{1}{2}<x<\frac{1}{2}+dx\right)$

Example Problems in Statistics

CCAR
Colorado Center for Astrodynamics Research

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

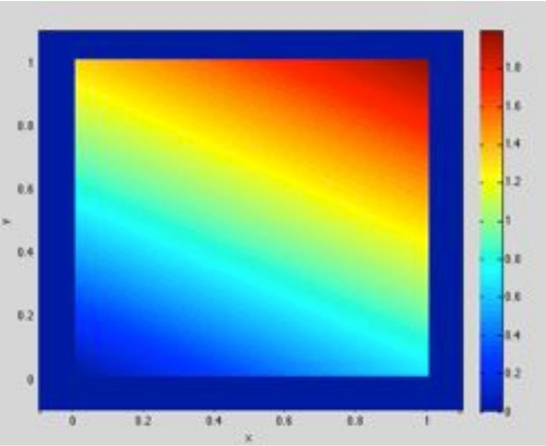
c) The probability that $x \le \frac{1}{2}$

d) $p\left(x+y \le \frac{1}{2}\right)$

e) $p\left(x+y = \frac{1}{2}\right)$

f) Whether x and y are independent

g) $p\left(0<y<\frac{1}{2} / \frac{1}{2}<x<\frac{1}{2}+dx\right)$



$$g(x/y)=\frac{f(x,y)}{h(y)}, \qquad h(y/x)=\frac{f(x,y)}{g(x)}$$

Example Problems in Statistics

CCAR
Colorado Center for Astrodynamics Research

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

g) $p\left(0<y<\frac{1}{2} / \frac{1}{2}<x<\frac{1}{2}+dx\right)$

$$h(y/x)=\frac{f(x,y)}{g(x)}$$

$$p\left(0<y<\frac{1}{2} / \frac{1}{2}<x<\frac{1}{2}+dx\right)=\int_0^{1/2}\frac{f\left(\frac{1}{2},y\right)}{g\left(\frac{1}{2}\right)}dy$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

g) $p\left(0<y<\frac{1}{2} / \frac{1}{2}<x<\frac{1}{2}+dx\right)$

$$h(y/x)=\frac{f(x,y)}{g(x)}$$

$$p\left(0<y<\tfrac{1}{2} / \tfrac{1}{2}<x<\tfrac{1}{2}+dx\right)=\int_0^{1/2}\frac{f\left(\frac{1}{2},y\right)}{g\left(\frac{1}{2}\right)}dy$$

$$=\int_0^{1/2}\frac{\frac{2}{3}\left(\frac{1}{2}+2y\right)}{\frac{2}{3}\left(\frac{1}{2}+1\right)}dy=\frac{\left[\frac{y}{2}+y^2\right]_0^{1/2}}{\frac{3}{2}}=\frac{1}{3}$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

Find:

a) k

b) The marginal density functions of x and y

c) The probability that $x \le \frac{1}{2}$

d) $p\left(x+y\le \frac{1}{2}\right)$

e) $p\left(x+y= \frac{1}{2}\right)$

f) Whether x and y are independent

g) $p\left(0<y<\frac{1}{2} / \frac{1}{2}<x<\frac{1}{2}+dx\right)$

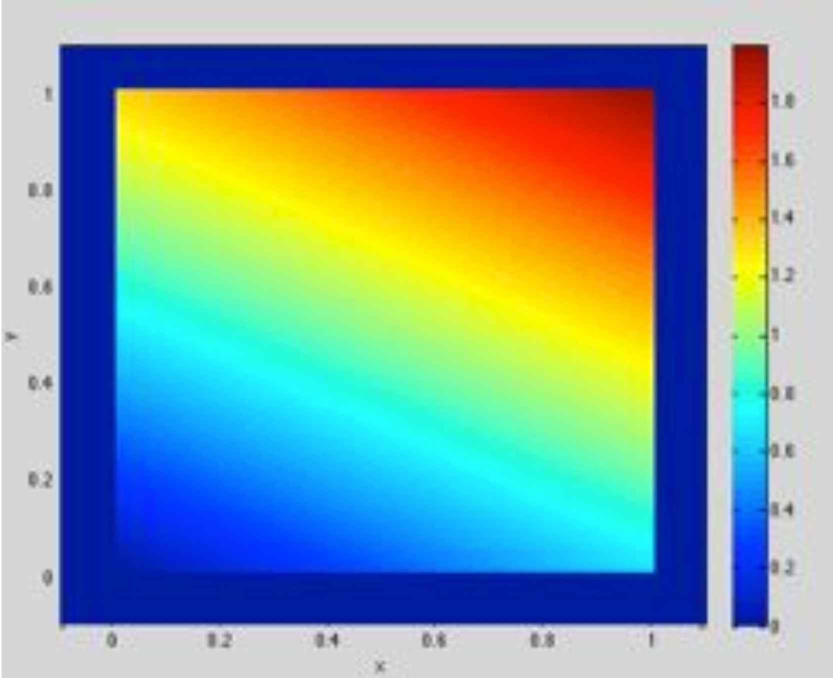
h) $p\left(0<y<\frac{1}{2} / x>\frac{1}{2}\right)$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

h) $p\left(0<y<\frac{1}{2} / x>\frac{1}{2}\right)$



Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

h) $p\left(0<y<\frac{1}{2}/x>\frac{1}{2}\right)$

$$=\frac{\int\limits_{y=0}^{1/2}\int\limits_{x=1/2}^{1} f(x,y)dxdy}{\int\limits_{1/2}^{1} g(x)dx}$$

Example Problems in Statistics

Given

$$f(x,y)=\begin{cases} k(x+2y) & 0\le x\le 1, 0\le y\le 1 \\ 0 & elsewhere \end{cases}$$

h) $p\left(0<y<\tfrac{1}{2} / x>\tfrac{1}{2}\right)$

$$=\frac{\int\limits_{y=0}^{1/2}\int\limits_{x=1/2}^{1} f(x,y)dxdy}{\int\limits_{1/2}^{1} g(x)dx}=\frac{\frac{2}{3}\int\limits_{0}^{1/2}\int\limits_{1/2}^{1}(x+2y)dxdy}{\frac{2}{3}\int\limits_{1/2}^{1}(x+1)dx}=\frac{5}{14}\approx 0.357$$

Example Problem continued

Determine the Variance-Covariance matrix for the example problem

$$f(x,y)=\frac{2}{3}(x+2y) \qquad 0\le x\le 1,\ \ 0\le y\le 1$$

$$=0 \qquad \text{elsewhere}$$

$$P=\begin{bmatrix} \sigma^2(x) & \mu_{11}(x,y) \\ \mu_{11}(x,y) & \sigma^2(y) \end{bmatrix}$$

$$\rho_{xy}=\frac{\mu_{11}}{\left[\sigma(x)\sigma(y)\right]}$$

Example Problem continued

Determine the Variance-Covariance matrix for the example problem

$$f(x,y)=\frac{2}{3}(x+2y) \qquad 0\le x\le 1,\ \ 0\le y\le 1$$

$$=0 \qquad \text{elsewhere}$$

We have shown that the marginal density functions are given by

$$g(x)=\frac{2}{3}[x+1]$$

$$h(y)=\frac{2}{3}\left[2y+\frac{1}{2}\right]$$

Example Problem continued

The elements of the variance-covariance matrix are computed below

$$E(x) = \lambda_{10} = \int_0^1 x g(x)\,dx = \frac{2}{3}\int_0^1 x(x+1)\,dx \qquad \lambda_{10} = E(x) = \frac{5}{9}$$

$$E(y) = \lambda_{01} = \int_0^1 y h(y)\,dy = \frac{2}{3}\int_0^1 y\left(2y + \tfrac{1}{2}\right)dy \qquad \lambda_{01} = E(y) = \frac{11}{18}$$

$$\sigma^2(x) = \mu_{20} = E\left[(x-\lambda_{10})^2\right] = E\left(x^2\right) - \lambda_{10}^2 = \frac{2}{3}\int_0^1 x^2(x+1)\,dx - \lambda_{10}^2$$

$$\mu_{20} = \frac{13}{162}$$

Example Problem continued

Cont.

$$\sigma^2(y) = \mu_{02} = E(y^2) - \lambda_{01}^2 = \frac{2}{3}\int_0^1 y^2\left(2y + \tfrac{1}{2}\right)dy - \lambda_{01}^2$$

$$\mu_{02} = \sigma^2(y) = \frac{23}{324}$$

$$\mu_{11} = E\left[(x-\lambda_{10})(y-\lambda_{01})\right] = E(xy) - \lambda_{10}\lambda_{01}$$

$$= \frac{2}{3}\int_0^1\int_0^1 xy(x+2y)\,dx\,dy - \lambda_{10}\lambda_{01}$$

$$\mu_{11} = -\frac{1}{162}$$

Example Problem continued

The variance-covariance matrix, P, is given by

$$P = \begin{bmatrix} \sigma^2(x) & \mu_{11}(x,y) \\ \mu_{11}(x,y) & \sigma^2(y) \end{bmatrix}$$

$$= \begin{bmatrix} 13/162 & -1/162 \\ -1/162 & 23/324 \end{bmatrix} = \begin{bmatrix} 0.0802 & -.00617 \\ -.00617 & 0.071 \end{bmatrix}$$

The correlation coefficient for random variables x and y is given by

$$\rho_{xy} = \frac{\mu_{11}}{\left[\sigma(x)\sigma(y)\right]} \quad = -0.082$$

Example Problem continued

The variance-covariance matrix, P, is given by

$$P = \begin{bmatrix} \sigma^2(x) & \mu_{11}(x,y) \\ \mu_{11}(x,y) & \sigma^2(y) \end{bmatrix}$$

$$= \begin{bmatrix} 13/162 & -1/162 \\ -1/162 & 23/324 \end{bmatrix} = \begin{bmatrix} 0.0802 & -.00617 \\ -.00617 & 0.071 \end{bmatrix}$$

The correlation coefficient for random variables x and y is given by

$$\rho_{xy} = \frac{\mu_{11}}{\left[\sigma(x)\sigma(y)\right]} \quad = -0.082$$

A common OD expression for the variance-covariance matrix has variances on the diagonal, covariance's in the upper triangle and correlation coefficients in the lower triangle.

Hence, $P = \begin{bmatrix} .0802 & -.00617 \\ -.082 & .071 \end{bmatrix}$

Statistical Problems

- Questions?

- One last quick subject:
  - Distributions with multiple variables that are Gaussian in any combination of those variables

Bivariate Normal Distribution

CCAR
Colorado Center for Astrodynamics Research

The *bivariate normal density function* is given by

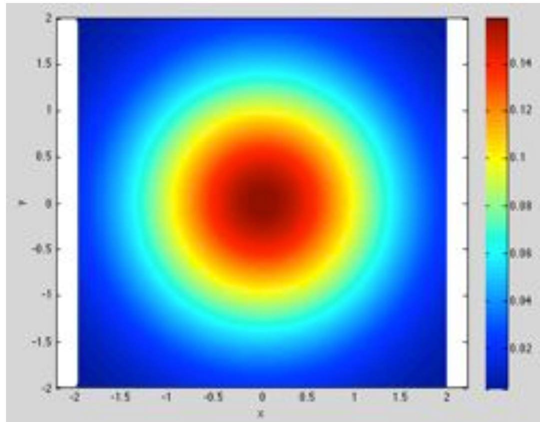
$$f(x,y) = \frac{1}{2\pi\,\sigma(x)\,\sigma(y)\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left[\frac{x-\lambda_{10}}{\sigma(x)}\right]^2 \right.\right. \qquad \begin{array}{l} -\infty < x < \infty \\ -\infty < y < \infty \end{array}$$

$$\left.\left. -2\rho\frac{[x-\lambda_{10}][y-\lambda_{01}]}{\sigma(x)\sigma(y)} + \left[\frac{y-\lambda_{01}}{\sigma(y)}\right]^2\right\}\right] \qquad \text{(A.17.1)}$$

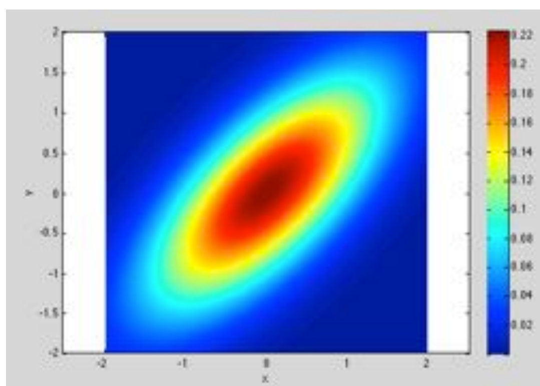
- Note that if $\rho = 0$, $f(x, y)$ may be factored into $f(x, y) = g(x)\,h(y)$.
- Hence, $\rho = 0$ is a sufficient condition for statistical independence of bivariate normal variables.

Bivariate Normal Distribution

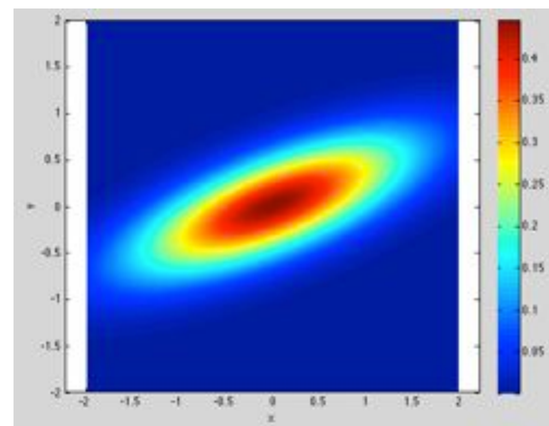
$\lambda(x) = 0.0, \ \sigma(x) = 1.0,$
$\lambda(y) = 0.0, \ \sigma(y) = 1.0$
$\rho = 0.0$



$\lambda(x) = 0.0, \ \sigma(x) = 1.0,$
$\lambda(y) = 0.0, \ \sigma(y) = 1.0$
$\rho = 0.7$



$\lambda(x) = 0.0, \ \sigma(x) = 1.0,$
$\lambda(y) = 0.0, \ \sigma(y) = 0.5$
$\rho = 0.7$



Marginal Density Function

- It can be shown that both X and Y in a bivariate normal density function have normal marginal density functions.

- By carrying out the integral,

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy \,. \tag{A.18.1}$$

It can be shown that

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma(x)} \exp\left[-\frac{1}{2}\left(\frac{x - \lambda_{10}}{\sigma(x)}\right)^2\right] \tag{A.18.2}$$

Similar results exist for the marginal density function of Y.

Conditional Density Function

- It can be shown that the conditional density functions for X and Y are also normal, i.e.,

$$h\left(y/x\right) = \frac{f\left(x, y\right)}{g\left(x\right)} = \frac{1}{\sigma\left(y\right)\sqrt{2\pi}\sqrt{1-\rho^2}}$$

$$\times \exp -\frac{1}{2}\left[\frac{y - \{\lambda_{01} + [\rho\sigma(y)/\sigma(x)][x - \lambda_{10}]\}}{\sigma\left(y\right)\sqrt{1-\rho^2}}\right]^2 . \qquad \text{(A.18.4)}$$

- Hence, the conditional density function of Y is normal with conditional mean

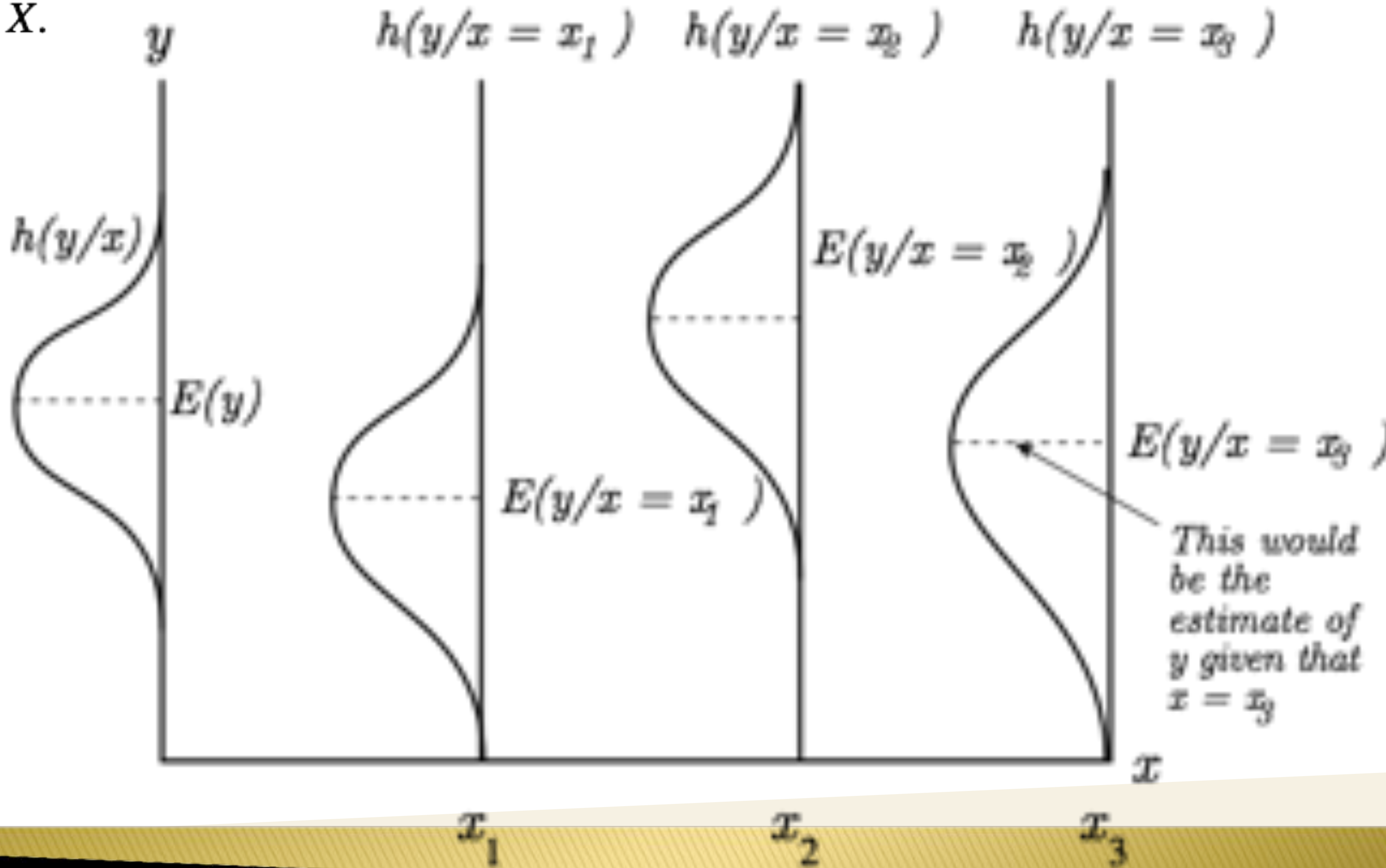
$$E\left(Y/x\right) = \lambda_{01} + \rho\frac{\sigma(y)}{\sigma(x)}[x - \lambda_{10}]$$

- and conditional standard deviation

$$\sigma(Y/x) = \sigma\left(y\right)\sqrt{1 - \rho^2} .$$

Conditional Density Function

The conditional mean as the estimate of Y is illustrated here for various realizations of X.



The Multivariate Normal Distribution

CCAR
Colorado Center for Astrodynamics Research

- For the multivariate case, consider a vector of random variables, e.g.,

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}.$$

- The *multivariate normal density function* is given by

$$f(x_1, x_2, \ldots, x_p) = \frac{1}{(2\pi)^{p/2}|V|^{1/2}} e^{-1/2\,(\mathbf{X}-\mathbf{\Lambda})^T V^{-1} (\mathbf{X}-\mathbf{\Lambda})}$$

$$-\infty < x_i < \infty \tag{A.19.1}$$

where

V	is the $p \times p$ variance-covariance matrix of the vector $\mathbf{X}$
$\lvert V \rvert$	is the determinant of V
$\mathbf{\Lambda}$	is a $p \times 1$ vector of mean values of X.

The Multivariate Normal Distribution

- The matrix V is defined as

$$V = E\{[\mathbf{X} - \mathbf{\Lambda}][\mathbf{X} - \mathbf{\Lambda}]^T\} \qquad \text{(A.19.2)}$$

- in terms of the correlation coefficient

$$\rho_{ij} = \frac{\mu_{ij}}{\sigma(x_i)\sigma(x_j)} \qquad \text{(A.19.3)}$$

$$V = \begin{bmatrix} \sigma^2(x_1) & \rho_{12}\sigma(x_1)\sigma(x_2) & \rho_{13}\sigma(x_1)\sigma(x_3) \cdots & \rho_{1p}\sigma(x_1)\sigma(x_p) \\ \rho_{12}\sigma(x_1)\sigma(x_2) & \sigma^2(x_2) & \rho_{23}\sigma(x_2)\sigma(x_3) \cdots & \rho_{2p}\sigma(x_2)\sigma(x_p) \\ \vdots & & & \\ \rho_{1p}\sigma(x_1)\sigma(x_p) & \rho_{2p}\sigma(x_2)\sigma(x_p) & \cdots \quad \cdots & \sigma^2(x_p) \end{bmatrix}.$$

Theorem 3: If the $p \times 1$ vector $\mathbf{X}$ is normally distributed with mean $\mathbf{\Lambda}$ and if the vector $\mathbf{X}$ is partitioned into two subvectors such that

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}, \mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 \\ \mathbf{\Lambda}_2 \end{bmatrix}, \text{ and } V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$$

then the conditional distribution of the $q \times 1$ vector $\mathbf{X}_1$, given that the vector $\mathbf{X}_2 = \mathbf{x}_2$, is the multivariate normal distribution with mean $\mathbf{\Lambda}_1 + V_{12} V_{22}^{-1} (x_2 - \mathbf{\Lambda}_2)$ and the covariance matrix $\left(V_{11} - V_{12} V_{22}^{-1} V_{21}\right)$.

Theorem 4: The covariance matrix of the conditional distribution of $\mathbf{X}_1$, given $\mathbf{X}_2 = \mathbf{x}_2$ does not depend on $\mathbf{x}_2$.

Proof: The proof of this is obvious from an examination of Eq. (A.19.5); that is, $V_{\mathbf{X}_1/\mathbf{X}_2=\mathbf{x}_2} = V_{11} - V_{12}V_{22}^{-1}V_{21}$. From Theorem 3, we also have

$$E(\mathbf{X}_1/\mathbf{X}_2 = \mathbf{x}_2) = \Lambda_1 + V_{12}V_{22}^{-1}(\mathbf{x}_2 - \Lambda_2). \tag{A.19.6}$$

If we were attempting to estimate $\mathbf{X}_1$, its mean value given by Eq. (A.19.6) would be a likely value to choose. Also, because the covariance of the conditional density function is independent of $\mathbf{x}_2$, we could generate the covariance without actually knowing the values of $\mathbf{X}_2$. This would allow us to perform an accuracy assessment of $\mathbf{X}_1$ without knowing the values of $\mathbf{X}_2$.

Final Statements

Colorado Center for Astrodynamics Research

- Homework 1 Graded
  - Comments included on D2L
  - Any questions, talk with us this week
- Homework 2 CAETE due Thursday
  - Graded soon after
- Homework 3 due Thursday
- Homework 4 out today
- Quiz available tomorrow at 1pm