Bi(Multi)variate Transformation of Variables

Daniel B. Rowe, Ph.D.

Professor
Department of Mathematical and Statistical Sciences

Outline

- **Bivariate Continuous Distributions**
  Joint PDF, Conditional PDF, Marginal PDF

- **Bivariate Transformation of Variables**
  Sum of RVs

Bivariate Continuous Distributions

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

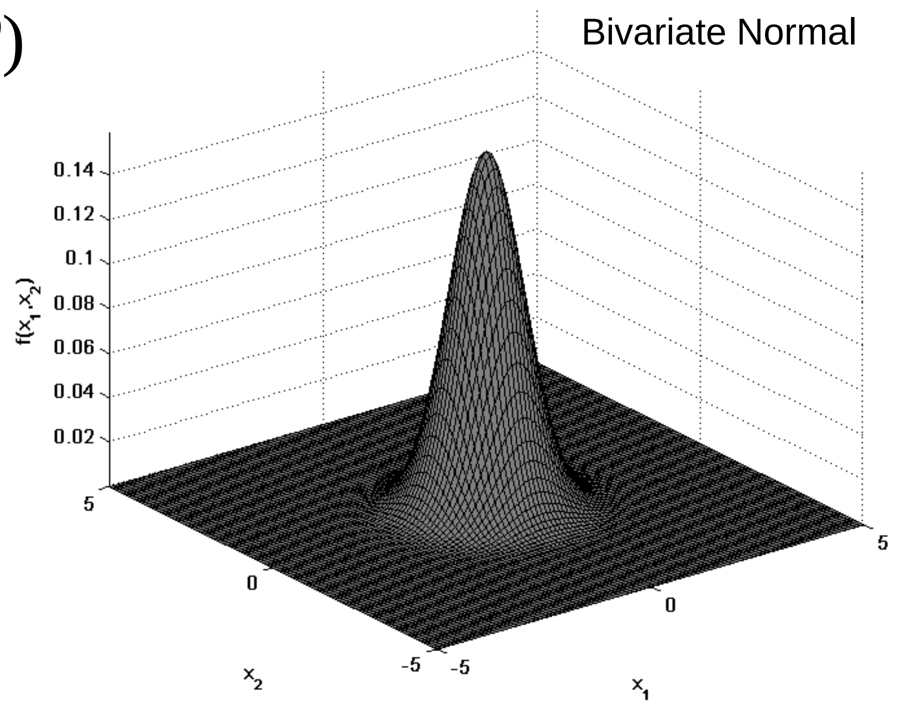
of two continuous random

variables (x_1, x_2) depending

upon parameters θ satisfies

1) $0 \le f(x_1, x_2 \mid \theta), \ \ \forall (x_1, x_2)$

2) $\iint\limits_{x_1 x_2} f(x_1, x_2 \mid \theta) dx_1 dx_2 = 1$.

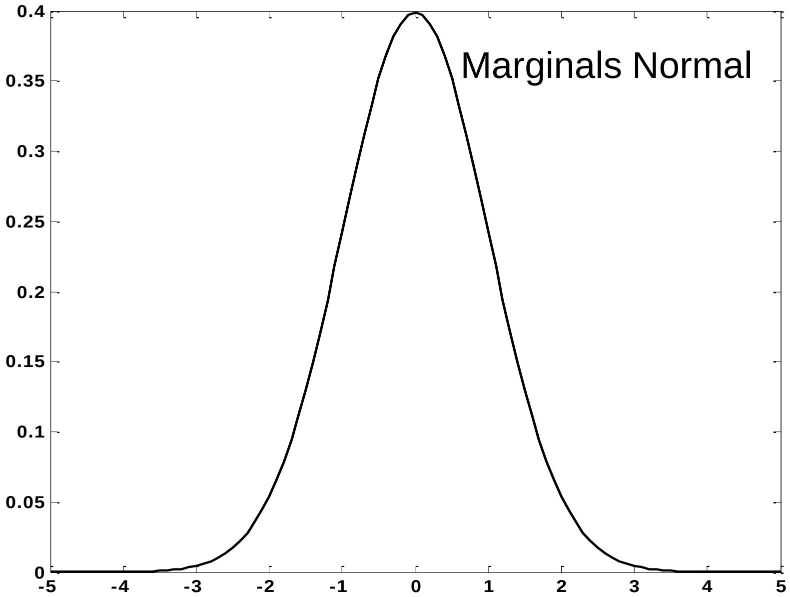


Bivariate Continuous Distributions

Marginal Distributions

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

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



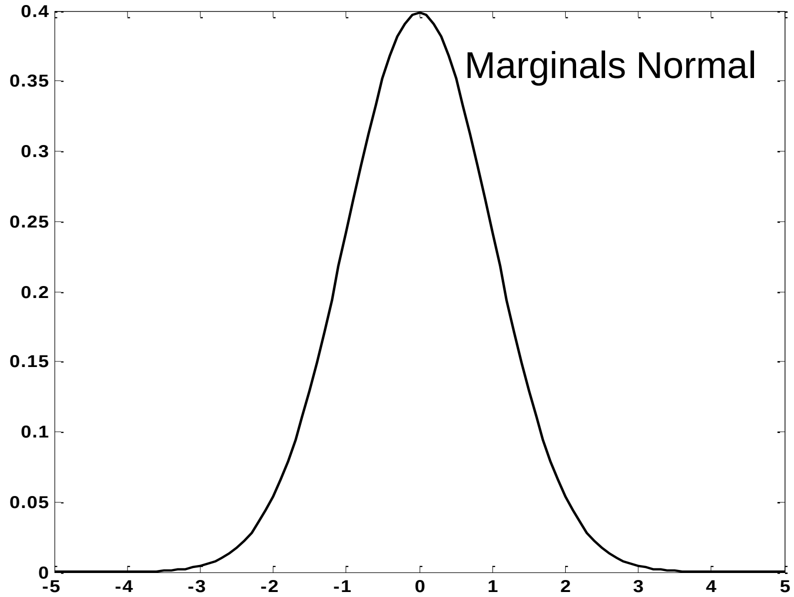
Marginal Expectations

$$E(g(X_1) \mid \theta) = \int_{x_1} g(x_1) f(x_1 \mid \theta) dx_1$$

$$E(g(X_2) \mid \theta) = \int_{x_2} g(x_2) f(x_2 \mid \theta) dx_2$$

Provided integral exists

Bivariate Continuous Distributions



Marginal Means

$$E(X_1 \mid \theta) = \int_{x_1} x_1 f(x_1 \mid \theta) dx_1 = \mu_1$$

$$E(X_2 \mid \theta) = \int_{x_2} x_2 f(x_2 \mid \theta) dx_2 = \mu_2$$

Marginal Variances

$$E([X_1 - E(X_1 \mid \theta)]^2 \mid \theta) = \int_{x_1} [x_1 - E(X_1 \mid \theta)]^2 f(x_1 \mid \theta) dx_1 = \sigma_1^2$$

$$E([X_2 - E(X_2 \mid \theta)]^2 \mid \theta) = \int_{x_2} [x_2 - E(X_2 \mid \theta)]^2 f(x_2 \mid \theta) dx_2 = \sigma_2^2$$

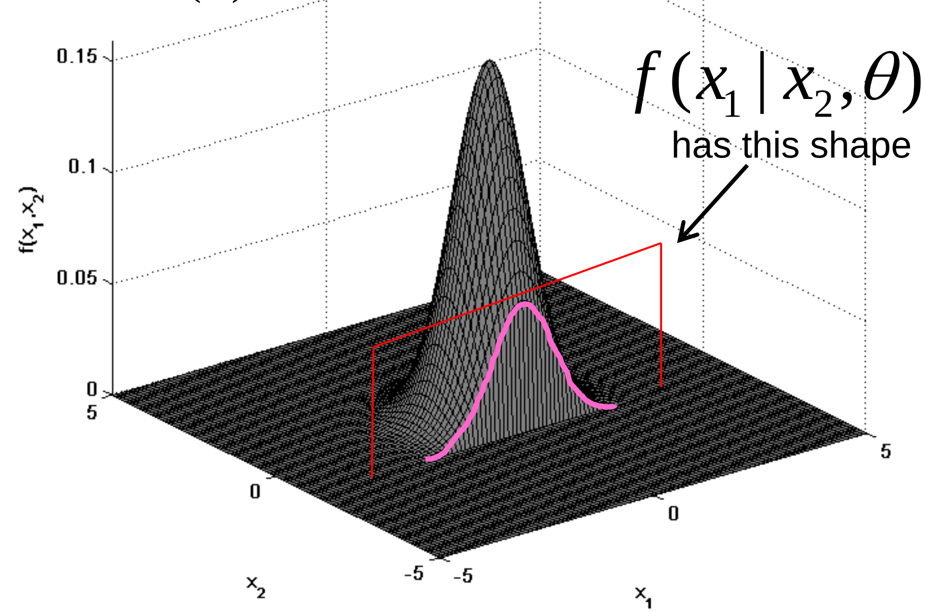
Bivariate Continuous Distributions

Conditional Distributions

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

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

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



Conditional Expectations

$$E(g(X_1) \mid X_2, \theta) = \int_{x_1} g(x_1) f(x_1 \mid x_2, \theta) dx_1$$

$$E(g(X_2) \mid X_1, \theta) = \int_{x_2} g(x_2) f(x_2 \mid x_1, \theta) dx_2$$

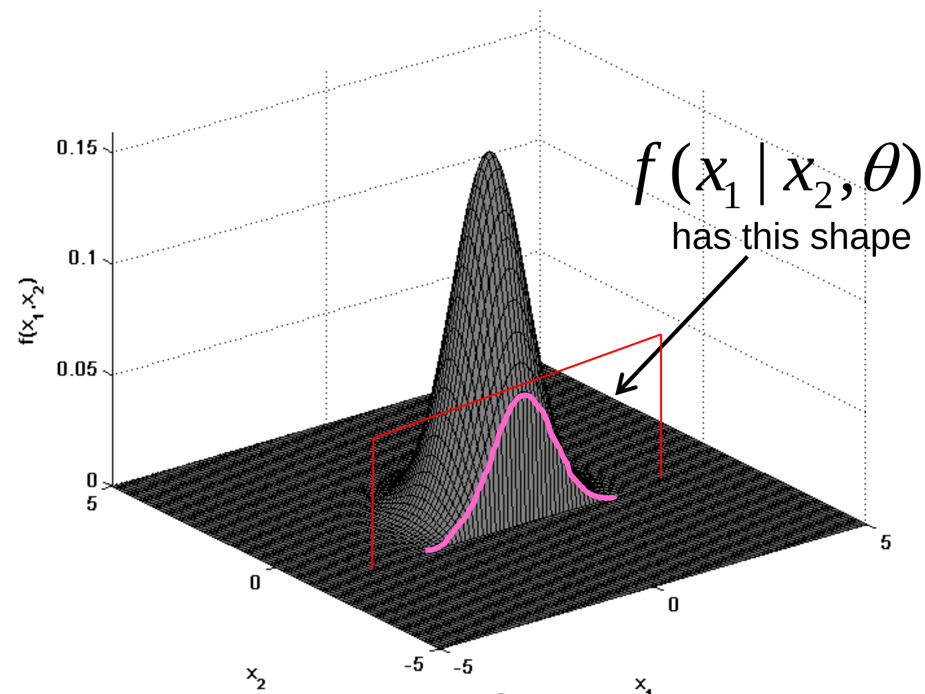
Provided integral exists

Bivariate Continuous Distributions

Conditional Means

$$E(X_1 \mid X_2, \theta) = \int_{x_1} x_1 f(x_1 \mid x_2, \theta) dx_1$$

$$E(X_2 \mid X_1, \theta) = \int_{x_2} x_2 f(x_2 \mid x_1, \theta) dx_2$$



Conditional Variances

$$E([X_1 - E(X_1 \mid X_2, \theta)]^2 \mid X_2, \theta) = \int_{x_1} [x_1 - E(X_1 \mid X_2, \theta)]^2 f(x_1 \mid x_2, \theta) dx_1$$

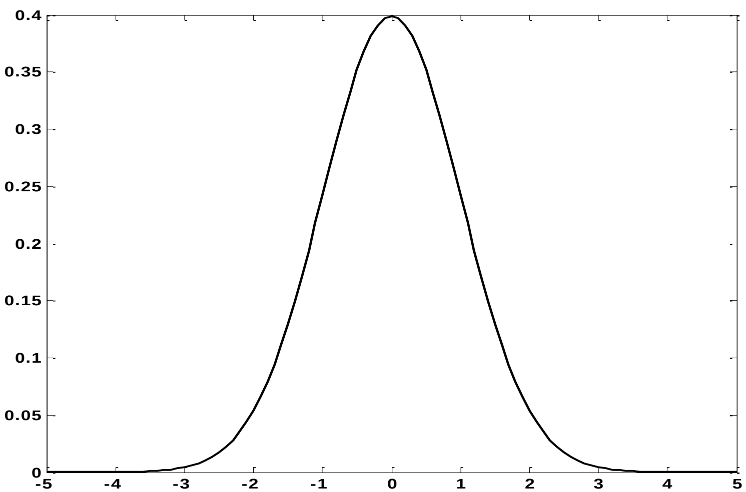
$$E([X_2 - E(X_2 \mid X_1, \theta)]^2 \mid X_2, \theta) = \int_{x_2} [x_2 - E(X_2 \mid X_1, \theta)]^2 f(x_2 \mid x_1, \theta) dx_2$$

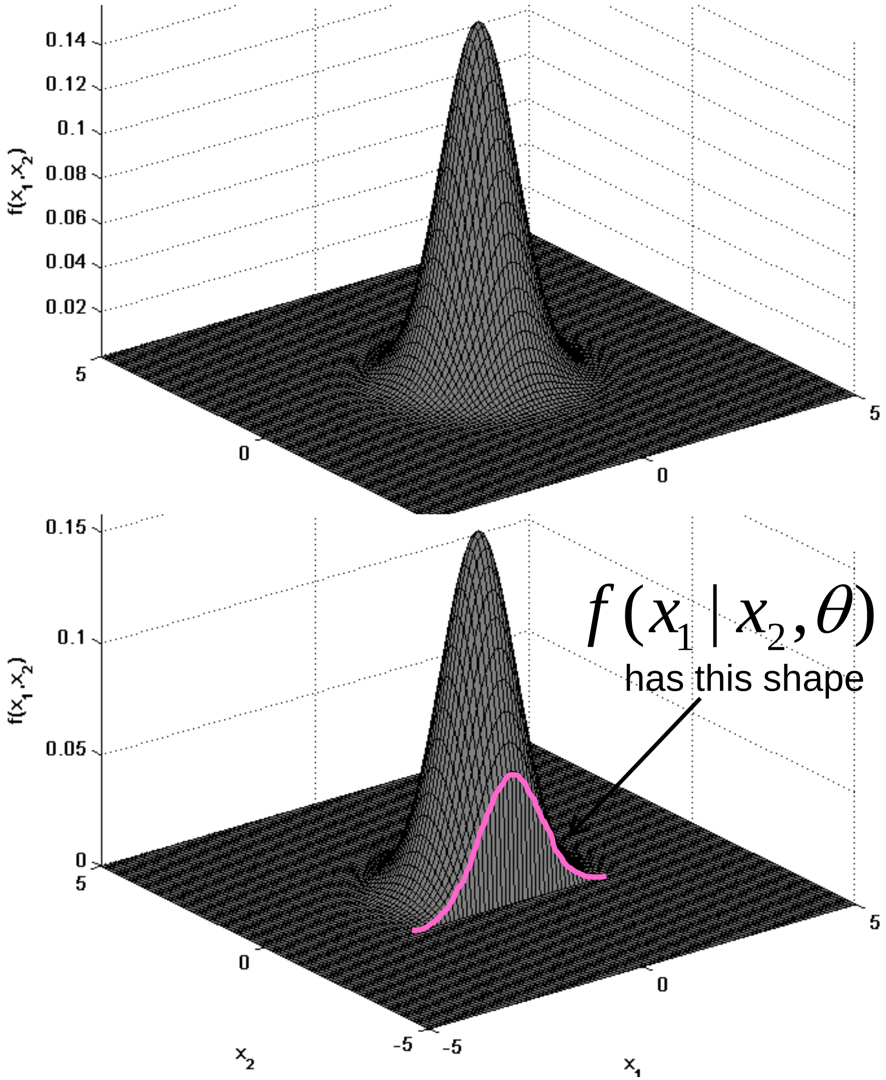
Bivariate Continuous Distributions

Marginal Distributions

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

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





Bivariate Continuous Distributions

Covariance

$$\text{cov}(X_1, X_2 | \theta) = \iint_{x_1 x_2} [x_1 - \underbrace{E(X_1 | \theta)}_{\mu_1}][x_2 - \underbrace{E(X_2 | \theta)}_{\mu_2}] f(x_1, x_2 | \theta) dx_1 dx_2$$

$$= \sigma_{12}$$

Correlation

$$\text{corr}(X_1, X_2 | \theta) = \frac{\text{cov}(X_1, X_2 | \theta)}{\sigma_1 \sigma_2}$$

Bivariate Continuous Distributions

Statistical Independence

Two random variables x_1 and x_2 are independent if and only if

$$f(x_1, x_2 \mid \theta) = f(x_1 \mid \theta_1) f(x_2 \mid \theta_2) .$$

If two random variables x_1 and x_2 are uncorrelated,

$$\text{corr}(X_1, X_2 \mid \theta) = 0$$

then they are not necessarily independent. (Only for normal).

Bivariate Change of Variable

Given two continuous random variables, (x_1,x_2)

with joint probability distribution function $f_{X_1,X_2}(x_1,x_2|\theta)$.

Let $\begin{pmatrix} y_1(x_1,x_2) \\ y_2(x_1,x_2) \end{pmatrix}$ be a transformation from (x_1,x_2) to (y_1,y_2)

with inverse transformation $\begin{pmatrix} x_1(y_1,y_2) \\ x_2(y_1,y_2) \end{pmatrix}$.

Bivariate Change of Variable

Then, the joint probability distribution function $f_{Y_1,Y_2}(y_1,y_2|\theta)$

of (y_1,y_2) can be found via

$$f_{Y_1,Y_2}(y_1,y_2\,|\,\theta)=f_{X_1,X_2}\left(x_1(y_1,y_2),x_2(y_1,y_2)\,|\,\theta\right)\times|\,J(x_1,x_2\to y_1,y_2)\,|$$

where $J(x_1,x_2\to y_1,y_2)=\begin{vmatrix}\dfrac{dx_1(y_1,y_2)}{dy_1} & \dfrac{dx_1(y_1,y_2)}{dy_2}\\ \dfrac{dx_2(y_1,y_2)}{dy_1} & \dfrac{dx_2(y_1,y_2)}{dy_2}\end{vmatrix}$.

$$|\,J(x_1,x_2\to y_1,y_2)\,|=1/\,|\,J(y_1,y_2\to x_1,x_2)\,|$$

Bivariate Change of Variable - Sum

Let x_1 have PDF $f_{X_1}(x_1|\theta)$ and x_2 have PDF $f_{X_2}(x_2|\theta)$,

then, the PDF of $y_1= x_1+x_2$ can be found via the

bivariate change of variable technique

$$f_{Y_1,Y_2}(y_1,y_2\,|\,\theta)= f_{X_1,X_2}(x_1(y_1,y_2),x_2(y_1,y_2)\,|\,\theta)\times|\,J(x_1,x_2\rightarrow y_1,y_2)\,|$$

with marginalization $f_{Y_1}(y_1\,|\,\theta)=\int_{y_2} f_{Y_1,Y_2}(y_1,y_2\,|\,\theta)dy_2$.

Bivariate Change of Variable - Sum

The joint PDF of (x_1,x_2) is

$$f_{X_1,X_2}(x_1,x_2 \mid \theta) = f_{X_1}(x_1 \mid \theta_1) f_{X_2}(x_2 \mid \theta_2) \,.$$ (need not be independent)

Let $y_1 = x_1+x_2$ and $y_2 = x_2$, then $x_1 = y_1 - y_2$ and $x_2 = y_2$.

$$J(x_1,x_2 \to y_1,y_2) = \begin{vmatrix} \dfrac{dx_1(y_1,y_2)}{dy_1} & \dfrac{dx_1(y_1,y_2)}{dy_2} \\ \dfrac{dx_2(y_1,y_2)}{dy_1} & \dfrac{dx_2(y_1,y_2)}{dy_2} \end{vmatrix} = 1$$

Bivariate Change of Variable - Sum

Example: Normal

Let x_1~normal(μ_1,σ_1^2) and x_2~normal(μ_2,σ_2^2), x_1&x_2 independent.
The joint PDF of (x_1,x_2) is

$$f(x_1,x_2\,|\,\mu_1,\sigma_1^2,\mu_2,\sigma_2^2)=\frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2}\frac{1}{\sigma_2\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2}.$$

With $x_1=y_1-y_2 \quad x_2=y_2 \quad J(x_1,x_2\to y_1,y_2)=1$

$$f_{Y_1,Y_2}(y_1,y_2\,|\,\theta)=\frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y_1-y_2-\mu_1}{\sigma_1}\right)^2}\frac{1}{\sigma_2\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y_2-\mu_2}{\sigma_2}\right)^2}\times 1$$

Bivariate Change of Variable - Sum

Rearranging leads to

$$f_{Y_1,Y_2}(y_1,y_2\,|\,\theta)=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}\left[\left(\frac{y_1-y_2-\mu_1}{\sigma_1}\right)^2+\left(\frac{y_2-\mu_2}{\sigma_2}\right)^2\right]}$$

Complete square in exponent to get

$$f_{Y_1,Y_2}(y_1,y_2\,|\,\theta)=\frac{1}{\tau\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y_2-\delta}{\tau}\right)^2}\frac{\tau}{\sigma_1\sigma_2\sqrt{2\pi}}e^{-\frac{1}{2}(\gamma-\tau^{-2}\delta^2)}$$

δ does not depend on y_2

$$\delta=\frac{\sigma_2^2(y_1-\mu_1)+\mu_1\sigma_2^2}{\sigma_1^2+\sigma_2^2}\qquad\tau^2=\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\qquad\gamma=\frac{(y_1-\mu_1)^2\sigma_2^2+\mu_2^2\sigma_1^2}{\sigma_1^2+\sigma_2^2}$$

Bivariate Change of Variable - Sum

Marginalizing leads to $f_{Y_1}(y_1 \mid \theta_1) = \int_{y_2} f_{Y_1,Y_2}(y_1, y_2 \mid \theta) dy_2$

$$f_{Y_1}(y_1 \mid \theta_1) = \frac{\tau}{\sigma_1 \sigma_2 \sqrt{2\pi}} e^{-\frac{1}{2}\left(\gamma - \tau^{-2}\delta^2\right)} \underbrace{\int_{y_2} \frac{1}{\tau\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_2-\delta}{\tau}\right)^2} dy_2}_{=1}$$

algebra leads to

$$f_{Y_1}(y_1 \mid \theta) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{[y_1 - (\mu_1 + \mu_2)]^2}{2(\sigma_1^2 + \sigma_2^2)}}$$

$$y_1 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

$$\tau^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

$$\delta = \frac{\sigma_2^2 (y_1 - \mu_1) + \mu_2 \sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

$$\gamma = \frac{(y_1 - \mu_1)^2 \sigma_2^2 + \mu_2^2 \sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

Bivariate Change of Variable - Sum

This change of variable technique can be repeated.

If x_3~normal(μ_3,σ_3^2), x_3 ind of x_1 & x_2, then if we let $y_3 = x_3 + y_1$ (don't forget $y_1 = x_1 + x_2$),

then we can find that $y_3 \sim N(\mu_1 + \mu_2 + \mu_3, \sigma_1^2 + \sigma_2^2 + \sigma_3^2)$

we can repeat the procedure to get $y_n \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$.

We can also find that $y = \frac{1}{n}\sum_{i=1}^{n} x_i \sim N\left(\frac{1}{n}\sum_{i=1}^{n} \mu_i, \frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2\right)$.

Homework 7:

1) Let x_1~normal(μ_1, σ_1^2), x_2~normal(μ_2, σ_2^2), x_1 and x_2 independent.

a) Use transformation to obtain the distribution of $y_1 = x_1 + x_2$.

b) Use transformation to obtain the distribution of $y_1 = x_1 - x_2$.

c) Use transformation to obtain the distribution of $y_1 = x_1 x_2$.

d) Use transformation to obtain the distribution of $y_1 = x_1 / x_2$.

In any of a)-d) you may need to constrain μ's and/or σ^2's.

(But try not to.)

e) Use numerical integration (rectangles) to get

$\mu_{y_1} = \int y_1 f(y_1) dy_1$ and $\sigma^2 = \int (y_1 - \mu_{y_1})^2 f(y_1) dy_1$ for a)-d).

f) Use numerical integration to get 50th and 99th percentiles.

Homework 7:

2) Let x_1~normal(5,4), x_2~normal(10,1), x_1 and x_2 independent.

a) For the transformation x_1 and x_2 to $y_1 = x_1 x_2$ and $y_2 = x_2$.

Use numerical integration with rectangular volumes to obtain the mean of y_2 from $f(y_1, y_2)$. Set up lower and upper bounds (a,b) for y_1 and (c,d) for y_2, then compute

$$\mu = \int_{y_1=-\infty}^{\infty} \int_{y_2=-\infty}^{\infty} y_1 f(y_1, y_2) dy$$

$$\sigma^2 = \int_{y_1=-\infty}^{\infty} \int_{y_2=-\infty}^{\infty} (y_1 - \mu)^2 f(y_1, y_2) dy_2 dy_1$$

⟵ can make 2 integrals

b) Repeat a) but for x_1 and x_2 to $y_1 = x_1/x_2$ and $y_2 = x_2$.

Homework 7:

3) Let x_1~normal(5,4), x_2~normal(10,1), x_1 and x_2 independent.
Generate 10^6 x_1's and 10^6 x_2's.

a) Let 10^6 new random variates be $y = x_1 + x_2$.

b) Let 10^6 new random variates be $y = x_1 - x_2$.

c) Let 10^6 new random variates be $y = x_1 x_2$.

d) Let 10^6 new random variates be $y = x_1 / x_2$.

e) For a)-d) generate histogram, means, variances
50th and 99th percentiles.

Compare results to 2). In any of a)-d) reconsider with any constraints you put on $\mu's$ and/or $\sigma^2 's$ in 1).

Homework 7:

4) Let $f(x \mid a,b,\alpha,\beta)=\alpha(x-\beta)^2$ for $x\in(a,b),\ a\in(-\infty,\infty),\ b\in(a,\infty)$

a) Sketch this probability function. Consider $a<\beta$ and $a>\beta$.

b) Find the normalizing constant α in terms of a and b.

c) Derive the mean of this distribution.

d) Derive the variance of this distribution.

e) Derive the median of this distribution.

f) What is the mode of this distribution.

g) Derive an expression for the CDF of this distribution.

Homework 7:

5) Let u~uniform(0,1).

a) Derive the distribution of $x=\sqrt[3]{3\frac{u}{\alpha}+(a-\beta)^3}+\beta, \quad x\in(a,b)$.

b) Generate 10^6 random u's and transform to x's.

use a=5, b=15, β=10.

c) Make a histogram of u's and x's.

d) Calculate mean, median, mode, and variance of x's.

e) Make ecdf for $F(x)$.

f) Repeat for a=7, β=10, b=17.

g) Compare your answers to their theoretical values.