Lesson 12

Derived Distributions

In real-life engineering applications, Systems are fed with random inputs, whose distributions are usually known.

We need to know the distribution of the outputs of various systems, given the distribution of their output. Previously we studied such derived distributions for discrete random variables.

In this lesson, we focus on more general cases, involving ~~of~~ continuous random variables.

Let us focus on $Y = g(X)$, where X is a continuous random variable and $g(\cdot)$ is a measurable function.

Every r.v. has a cdf.

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(g(X) \le y)$$

$$= \mathbb{P}(X \in B_y) = \mathbb{P}_X(B_y)$$

where

$$B_y = \{x \in \mathbb{R} \mid g(x) \le y\}$$

B_y is a Borel Set (why?)

If the pdf of X is f_X then

$$F_Y(y) = \int_{B_y} f_X(x)\,dx$$

And the pdf of Y is:

$$f_Y(y) = \frac{d F_Y(y)}{dy}$$

Another equivalent approach would be using the probability law of Y:

$\forall B \in \mathcal{B}(\mathbb{R})$

$$\mathbb{P}_Y(B) = \mathbb{P}(Y \in B) = \mathbb{P}(g(X) \in B)$$

$$= \mathbb{P}(X \in \overleftarrow{g}(B)) = \mathbb{P}_X(\overleftarrow{g}(B))$$

Example: Assume ~~X~~ $X \sim$ Laplace(1).

$Y = |X|$. What is the cdf of Y?

What is the pdf of Y?

Example: Find the cdf and density of $Y = g(X)$ if $X \sim U[0, 4]$ and

$$g(x) = \begin{cases} x & 0 \le x < 1 \\ 1 & 1 \le x < 2 \\ 3-x & 2 \le x < 3 \\ 0 & \text{otherwise} \end{cases}$$

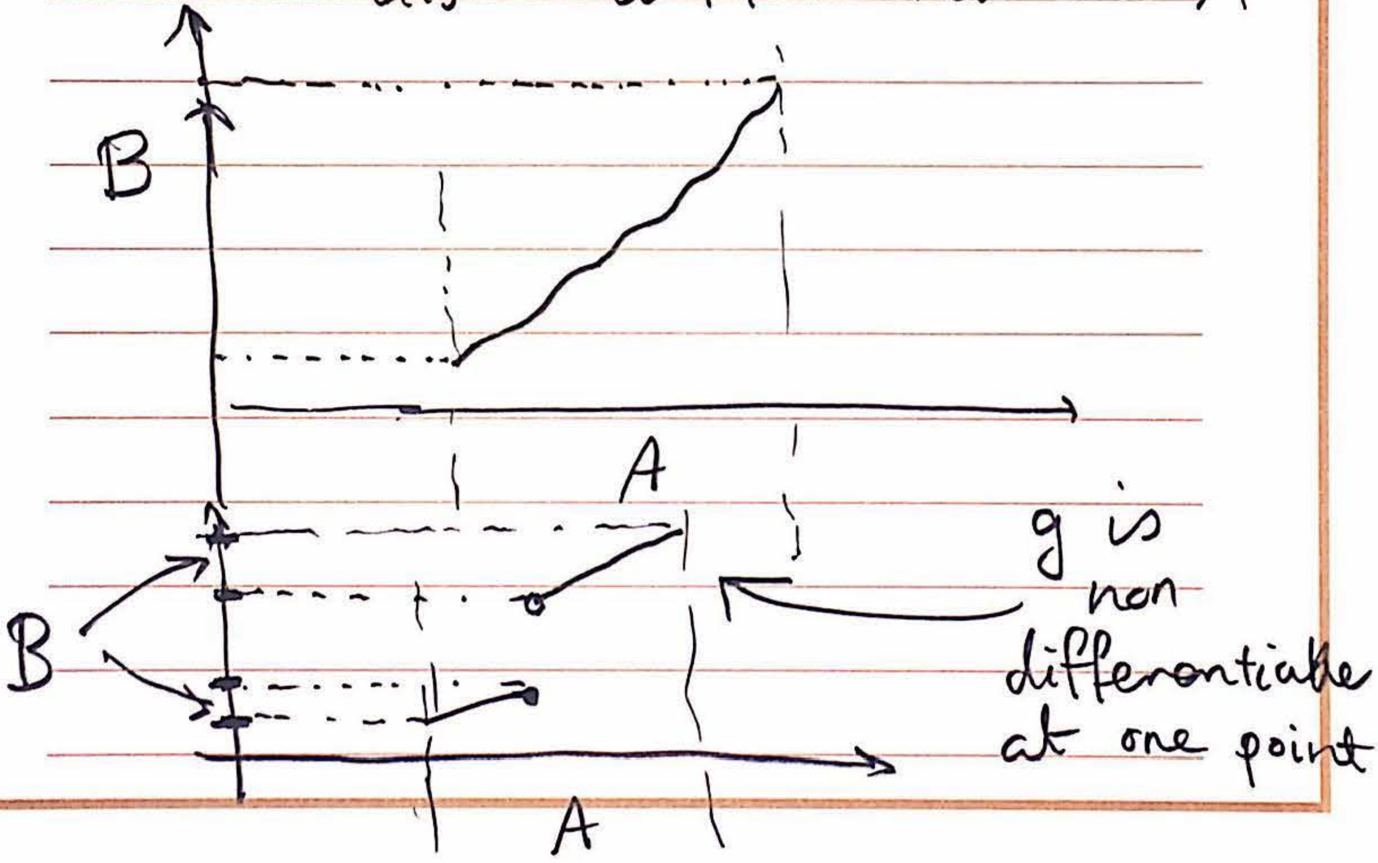
Example: Let $Y = g(X)$, where X is a continuous r.v. with known pdf. Assume $g(x) = x^2$. Find the pdf and cdf of $Y = g(X)$.

Exercise: Assume that X is a continuous r.v. and $X \geq 0$ with probability 1. Also assume that $Y = \exp(X^4)$. Find the pdf and cdf of Y.

Monotonic Functions

Assume that the range of $\vec{X}$ contains an open interval, A, i.e. $A \subseteq \vec{X}(\Omega)$. Also assume that g is strictly monotonic on A

and is also differentiable on A



Let B be the set of all values of $g(x)$, when $x \in A$, i.e. $B = \vec{g}(A)$. First let's assume g is increasing. Since g is monotonic, it has an inverse function. Let g^{-1}

be the inverse function of g.

Then $g(g^{-1}(y)) =$

$\forall y \in B$

$\forall y \in B$,

$F_Y(y) = \mathbb{P}(Y \leq y)$

Then

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$

$$=$$

Using the chain rule:

$f_Y(y) =$

Next, one needs to calculate

$\frac{d\, g^{-1}(y)}{dy}$. Because $g(g^{-1}(y)) = y$,

Using the chain rule:

substitution of the expression for $\frac{dg^{-1}(y)}{dy}$ in the formula

When g is monotonic and decreasing,

$F_Y(y) = \mathbb{P}(g(X) \leq y)$

$= \mathbb{P}(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y))$

therefore

$f_Y(y) =$

Combining these two cases,

one has

$$f_Y(y) = f_X(g^{-1}(y)) \frac{1}{|g'(g^{-1}(y))|}$$

$y \in B$

Linear Functions

Consider the special case

$a \neq 0$, $g(x) = ax + b$. $Y = aX + b$.

Then:

Example: A linear function of a normal r.v.:

Assume that $X \sim N(\mu, \sigma^2)$.

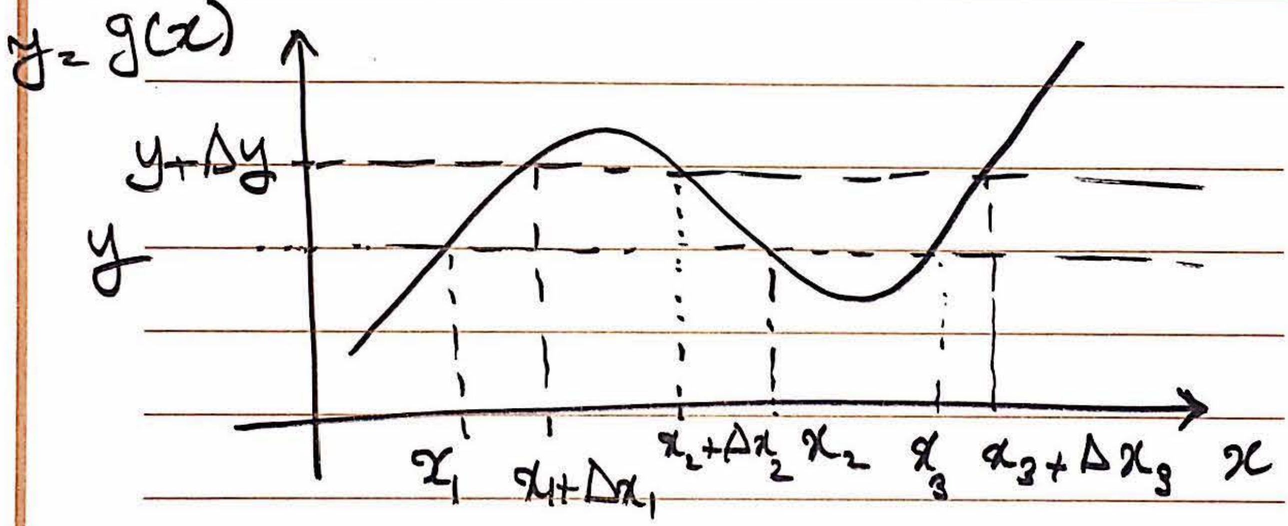
Find the pdf of $Y = aX + b, a \neq 0$

Example: Assume that X is distributed according to a standard Cauchy. What is the pdf of $Y = g(X) = X^3 + 1$.

Non-monotonic Functions

Usually, we deal with non-monotonic functions of a r.v., that may not even be continuous.

Consider the following function:



Assume that one wishes to determine the pdf of $Y = g(X)$

Observe that if

$$y \leq Y \leq y + \Delta y$$

then for X

where x_1, x_2, x_3 are roots

of the equation $y = g(x)$

In general, for a function

$g(x)$ that is locally differentiable

and piece-wise monotonic

if the event $\{y \leq Y \leq y + \Delta y\}$ occurs,

then

$$X \in [x_i, x_i + \Delta x_i]$$

if $g(x)$ is increasing around x_i

and $X \in [x_i + \Delta x_i, x_i]$

if $g(x)$ is decreasing around x_i.

Therefore

$$P(y \leq Y \leq y + \Delta y) = \sum_{i} P(x_i \leq X \leq x_i + \Delta x_i)$$

$$+ \sum_{j} P(x_j + \Delta x_j \leq X \leq x_j)$$

where the sum is broken

according to the regions in which $g(x)$ is decreasing or increasing. Using pdfs and cdfs

$$f_Y(y)\,|\Delta y| = \sum_k$$

Example: Assume that $X \sim Laplace(1)$ and $Y = |X|$. Find the pdf of Y.

Multivariable Functions

(Functions of Random Vectors)

Suppose $\underline{X} = [X_1, \dots, X_n]^T$ is a random vector, and that X_i's are jointly continuous with joint pdf $f_{\underline{X}}(\underline{x}) = f_{\underline{X}}(x_1, \dots, x_n)$

Consider a function $g : \mathbb{R}^n \to \mathbb{R}^n$ and the random vector $\underline{Y} = g(\underline{X})$

$\underline{Y} = [Y_1, Y_2, \dots, Y_n]$.

Let g_i be the i^{th} component of

g, so that $Y_i = g_i(\underline{X}) = g_i(X_1, \dots, X_n)$.

Example: Let $g: \mathbb{R}^2 \to \mathbb{R}^2$

$$g(x_1, x_2) = \begin{bmatrix} g_1(x_1, x_2) \\ g_2(x_1, x_2) \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$g\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1^2 + x_2^2 \\ x_1 x_2 \end{bmatrix}$$

$y_1 = g_1(x_1, x_2) = x_1^2 + x_2^2$

$y_2 = g_2(x_1, x_2) = x_1 x_2$

We are interested in finding the joint pdf of $Y_1, \dots, Y_n$ based on the joint pdf of $X_1 \dots X_n$.

Linear Functions

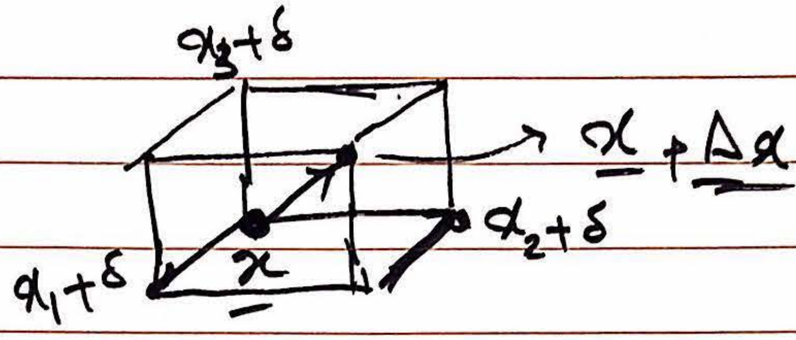
Assume that

$$g(\underline{x}) = \underline{G}\,\underline{x}, \text{ where } \underline{G} \in \mathbb{R}^{n \times n}$$

Assume that $A \subseteq \mathbb{R}^n$ is an open subset of $\mathbb{R}^n$. Fix

$\underline{x} \in A$ and $\Delta \underline{x} = [\delta, \dots, \delta]^T$, where δ is small, and

$\underline{x} + \Delta \underline{x}$ represents a hypercube

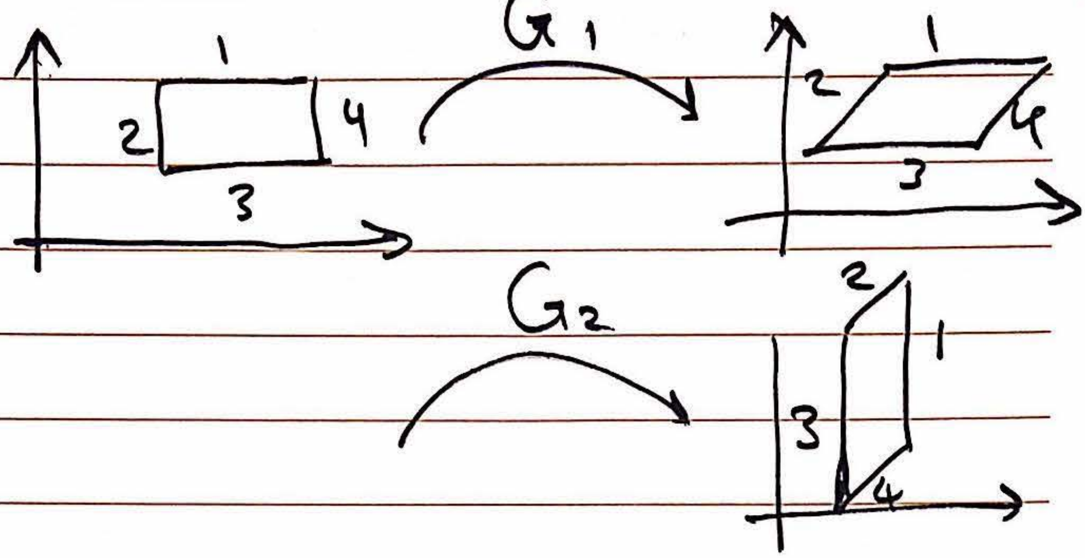


Assume that $\underline{\Delta x}$ is small enough to keep the hypercube inside A. Let's

call the set of points inside the hypercube, C.

A linear mapping "preserves" parallel lines, but not necessarily angles:

Ex:



The image of C under G is "parallelpiped," i.e. all of its faces are parallelograms.

Let's call the image of C $\vec{g}(C)$, P.

$$\mathbb{P}(\underline{X} \in C) = \mathbb{P}(g(\underline{X}) \in \vec{g}(C))$$
$$= \mathbb{P}(g(\underline{X}) \in P)$$

On the other hand C and

$P = \vec{g}(C)$ are Borel subsets of $\mathbb{R}^n$, therefore

$$\mathbb{P}(\underline{X} \in C) = \int_C$$

$\approx$

and

$$\mathbb{P}(g(\underline{X}) \in P) = \mathbb{P}(\underline{Y} \in P)$$

$$= \int_P$$

$$\simeq f_{\underline{Y}}(\underline{y}) \cdot Vol(P)$$

From analytic geometry, we know that the Volume of P is

$$Vol(P) = |\det(G)|\ Vol(C)$$

$$= |\det(G)|\ \delta^n$$

because

$$\mathbb{P}(\underline{X} \in C) = \mathbb{P}(\vec{g}(\underline{X}) \in \vec{g}(C))$$

$$= \mathbb{P}(Y \in P)$$

$$f_{\underline{X}}(\underline{x})\, \delta^n = f_{\underline{Y}}(\underline{y})\, |\det(G)|\, \delta^n$$

$$\Rightarrow f_{\underline{X}}(\underline{x}) = f_{\underline{Y}}(\underline{y})\, |\det(G)|$$

$$\Rightarrow f_{\underline{Y}}(\underline{y}) = \frac{f_{\underline{X}}(\underline{x})}{|\det(G)|}$$

when $\det(G) \neq 0$.

when $\det(G) \neq 0$, G is invertible,

So $\underline{y} = G\underline{x} \Rightarrow \underline{x} = G^{-1}\underline{y}$

$$\Rightarrow f_{\underline{y}}(\underline{y}) = \frac{f_{\underline{X}}(G^{-1}\underline{y})}{|\det(G)|}$$

$$= f_X(G^{-1}\underline{y})\,|\det(G^{-1})|$$

What happens if G is singular?

If G is singular, the linear transformation $\underline{y} = G\underline{x}$ cannot be onto.

{Reason: Fundamental Theorem of Linear Algebra

$$\underbrace{\text{Rank}}_{\leq n} + \underbrace{\text{Nullity}}_{>0} = n$$

when nullity >0, a subspace of $\mathbb{R}^n$ is mapped into $\underline{0}$,

so the linear transformation cannot cover all $\mathbb{R}^n$.

Example: Assume that $G = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

and $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$

$$\underline{Y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 + X_2 \\ X_1 + X_2 \end{bmatrix}$$

$$= \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

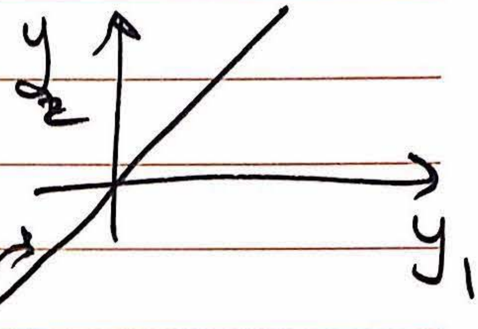
Obviously $Y_1 = Y_2$, therefore Y_1 and Y_2 are <u>not</u> jointly continuous

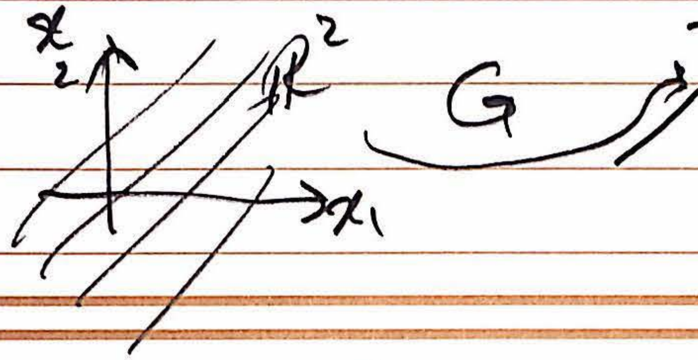
G maps $\mathbb{R}^2$ on a line:

$$\underline{x} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \quad \underline{y} = G\underline{x} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + x_2 \\ x_1 + x_2 \end{bmatrix}$$

$$\Rightarrow y_1 = y_2$$





In general, a singular G maps $\mathbb{R}^n$ to a subspace of $\mathbb{R}^n$ with a - dimension lower than n

Monotonic Functions

Let us generalize the previous discussion to a case where g is continuously differentiable at x on some open

Set $A \subset \mathbb{R}^n$. Let $B = \vec{g}(A)$ be the image of A under g and also assume that g is invertible on B, i.e. g^{-1} is a well-defined function, i.e.

$\forall \underline{y} \in \hat{B}, \exists! \ \underline{x} \in \mathbb{R}^n$ s.t.

$g(\underline{x}) = \underline{y}$

Since we chose δ very small,

the image of the hypercube C

is approximately parallelpiped,

since in a small neighborhood

of $\underline{x}$, $g(\underline{x})$ can be approximated

by a linear function using

Taylor Series:

$$g(\underline{x} + \underline{\Delta x}) \simeq g(\underline{x}) + G(\underline{x}) \, \underline{\Delta x} \quad \forall \underline{x} \in C$$

where $G(\underline{x})$ is the Jacobian Matrix of g at $\underline{x}$

$$G(\underline{x}) = \left(\frac{\partial g}{\partial \underline{x}}\right)(\underline{x})$$

$$G_{ij} = \frac{\partial g_i(\underline{x})}{\partial x_j}$$

Again, the volume of $P = \vec{g}(C)$ can be approximated as

$$Vol(P) \approx |\det(G(\underline{x}))| \, \delta^n$$

Therefore:

$$f_{\underline{Y}}(\underline{y}) = \frac{f_{\underline{X}}(g^{-1}(\underline{y}))}{\left|\det G(g^{-1}(\underline{y}))\right|}$$

Recall that:

$$g(g^{-1}(\underline{y})) = \underline{y}$$

$$\frac{\partial}{\partial \underline{y}}\left(g(g^{-1}(\underline{y}))\right) = \frac{\partial \underline{y}}{\partial \underline{y}} = I_{n\times n}$$

$$\Rightarrow \frac{\partial g^{-1}(\underline{y})}{\partial \underline{y}} \underbrace{g'(g^{-1}(\underline{y}))}_{G(g^{-1}(\underline{y}))} = I_{n\times n}$$

$$\Rightarrow \left|\det \frac{\partial g^{-1}(\underline{y})}{\partial \underline{y}}\right| \left|\det G(g^{-1}(\underline{y}))\right| = 1$$

$$\Rightarrow \left| \det \frac{\partial g^{-1}(\underline{y})}{\partial \underline{y}} \right| = \frac{1}{\left| \det G(g^{-1}(\underline{y})) \right|}$$

Denote $\frac{\partial g^{-1}(\underline{y})}{\partial \underline{y}} = H(\underline{y})$

then

$$f_{\underline{Y}}(\underline{y}) = \frac{f_{\underline{X}}(g^{-1}(\underline{y}))}{\left| \det(G(g^{-1}(\underline{y}))) \right|}$$

$$= f_{\underline{X}}(g^{-1}(\underline{y})) \left| \det(H(\underline{y})) \right|$$

Example: Suppose X_1 and X_2 have the joint pdf

$$f_{X_1 X_2}(x_1, x_2) = \begin{cases} 2 & 0 < x_1 < x_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Compute the joint pdf of $Y_1 = \frac{X_1}{X_2}$ and $Y_2 = 2X_2$

Exercise: X_1 and X_2 are jointly continuous with joint pdf $f_{X_1,X_2}(x_1,x_2)$. Find the joint pdf of $Y_1 = \frac{X_1}{X_1+X_2}$ and $Y_2 = 1/X_2$.

Non-invertible functions

Similar to single-variable case

$$f_{\underline{Y}}(\underline{y}) = \sum_{k} \frac{f_{\underline{X}}(\underline{x})}{\left|\det(G(\underline{x}))\right|}\Bigg|_{\underline{x}=\underline{x}_k}$$

where $\underline{x}_k$ is the root of $g(\underline{x}) = \underline{y}$ and $G(\underline{x}) = \partial g(\underline{x}) / \partial \underline{x}$

A Single Function of Multiple Random Variables

Suppose that $\underline{X} = [X_1, \ldots, X_n]^T$ is a vector of jointly continuous random variables with joint pdf $f_{\underline{X}}(\underline{x})$.

Consider a function $g_1 : \mathbb{R}^n \to \mathbb{R}$

We are interested in the pdf of $Y_1 = g_1(\underline{X})$. For example

$$g(x_1, x_2, x_3) = \frac{x_1 x_2}{x_3}$$

Basically, one can use the definition of cdf and pdf to calculate $f_{Y_1}(y)$:

$$F_{Y_1}(y) = \mathbb{P}\left(g(\underline{X}) \le y\right)$$

$$= \int_{\{\underline{x} \mid g(\underline{x}) \le y\}} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

where $d\underline{x} = dx_1 \, dx_2 \cdots dx_n$

and

$$f_{Y_1}(y) = \frac{dF_{Y_1}(y)}{dy}$$

Another interesting way of calculating $f_{Y_1}(y)$ is to use formulas for multivariable functions.

Let's assume $g_2, \dots, g_n : \mathbb{R}^n \to \mathbb{R}$ are functions and $Y_i = g_i(\underline{X})$ $i \in \{1, \dots, n\}$

If the function

$$g(\underline{x}) = (g_1(\underline{x}), g_2(\underline{x}), \dots, g_n(\underline{x}))$$

is invertible, we can find the joint pdf of Y_i's using previously discussed methods. Then, the marginal pdf of Y_1 is easily found by integrating $y_2, \ldots, y_n$ out, from the joint pdf.

A very simple choice for

g_2 to g_n is the following

$$g(\underline{x}) = (g_1(\underline{x}), x_2, \ldots, x_n)$$

i.e. $y_i = g_i(\underline{x}) = x_i \quad \forall i \geq 2.$

Since $g^{-1}(\underline{y})$ is assumed to exist

$$g^{-1}(\underline{y}) = \begin{bmatrix} h_1(\underline{y}) \\ h_2(\underline{y}) \\ \vdots \\ h_n(\underline{y}) \end{bmatrix}$$

but $\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} g_1(\underline{x}) \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$\Rightarrow$ $y_2 = x_2$ $\Rightarrow$ $x_2 = h_2(y) = y_2$

$y_3 = x_3$ $\qquad x_3 = h_3(y) = y_3$

$\vdots \qquad \vdots$

$y_n = x_n \qquad x_n = h_n(y) = y_n$

$$\Rightarrow g^{-1}(\underline{y}) = \begin{bmatrix} h_1(\underline{y}) \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$H(\underline{y}) = \frac{\partial g^{-1}(\underline{y})}{\partial \underline{y}}$$

$$= \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}$$

$$\Rightarrow \left| \det H(\underline{y}) \right| =$$

Therefore:

$$f_{\underline{Y}}(\underline{y}) = f_{\underline{X}}$$

and

$$f_{Y_1}(y_1) = \int f_{\underline{X}}\left(h_1(\underline{y}), y_2, \dots, y_n\right) \left|\frac{\partial h_1}{\partial y_1}\right| dy_2\, dy_3 \cdots dy_n$$

Example: Let X_1 and X_2 be positive jointly continuous r.v.'s and suppose we wish to derive the pdf of

$$Y_1 = g(X_1, X_2) = X_1 X_2$$

Introduce $Y_2 = X_2$, and derive $f_{Y_1}(y_1)$. In particular solve the problem when $X_1 \sim U(0,1)$ and $X_2 \sim U(0,1)$. and are independent.
Also, solve the problem directly

using multiple integrals

Order Statistics

An important set of many to one functions $g: \mathbb{R}^n \to \mathbb{R}$ are order statistics. Assume that $X_1, X_2, \ldots, X_n$ are i.i.d and jointly

Continuous. Let's call an ordered version of X_i's, $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$. Clearly, $X_{(1)} = \min_i X_i$ and $X_{(n)} = \max_j X_j$.

First, let's calculate the pdf of

$Y = \min_i (X_i)$ and $Z = \max_j (X_j)$.

$F_Y(y) = \mathbb{P}(Y \leq y)$

$$F_X(x) = \mathbb{P}(X \le x)$$

Now, let's calculate the pdf of $X_{(k)}$.

$$\mathbb{P}(x \leq X_{(k)} \leq x + \Delta x) = f_{X_{(k)}} \Delta x$$

for small Δx

Exercise: Show that the joint pdf of $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ is in general

$$f_{X_{(1)},\ldots,X_{(n)}}(x_1,\ldots,x_n) = \begin{cases} n!\, f(x_1, x_2, \ldots, x_n) & x_1 < x_2 < \cdots < x_n \\ 0 & \text{otherwise} \end{cases}$$

and if $X_1 \ldots X_n$ are i.i.d

$$f_{X_{(1)},\ldots,X_{(n)}}(x_1,\ldots,x_n) = \begin{cases} n! \prod_{i=1}^{n} f_X(x_i) & x_1 < x_2 < \cdots < x_n \\ 0 & \text{otherwise.} \end{cases}$$

Sum of Independent Random Variables

Previously, we observed that if X and Y are independent discrete r.v.'s and $Z = X + Y$,

$$P_Z = P_X * P_Y$$

For continuous r.v.'s: $Z = X + Y$

$$F_Z(z) = P(X + Y \leq z) = \iint_{(x,y) | x+y \leq z} f_{X,Y}(x,y)\,dx\,dy$$

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \frac{d}{dz} \mathbb{P}(X+Y \le z)$$

$$=$$

Exercise: Assume that $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ are independent.

Show that $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

By induction

If $X_1, X_2, \dots, X_n$ are independent r.v.'s and $Z = X_1 + X_2 + \dots + X_n$

$$f_Z = f_{X_1} * f_{X_2} * \dots * f_{X_n}$$

Computer Exercise:

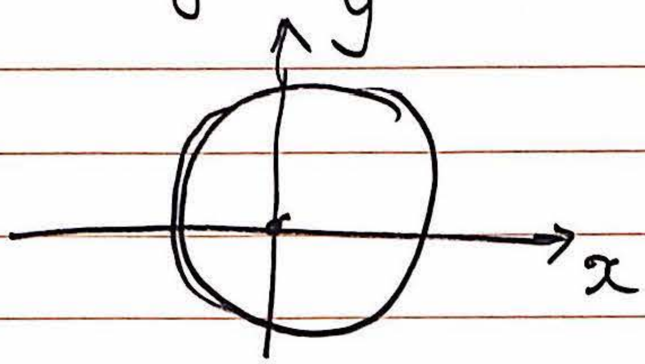
Assume that $X_1, \dots, X_n$ are i.i.d $U(0,1)$ r.v.'s. Assume that $n=5$, $n=10$, $n=30$, $n=100$.

Plot the pdf of $Z = X_1 + \dots + X_n$.

Also, repeat the exercise by sampling from $X_1, \ldots, X_n$ and creating a distribution for $Z = X_1 + \cdots + X_n$, using histograms or Kernel Density Estimation

Some Derived Distributions

Assume that a shooter shoots at a target.



Assume that the X coordinate of the shot is a $N(0,1)$ r.v. and also the Y coordinate of the shot is a $N(0,1)$ r.v. and X, Y are independent.

What are the distributions of R and θ, the distance of the shot from the origin and its angle with the x axis?

Let's define $(r, \theta) = g(x, y)$

We need the joint distribution of $(R, \Theta) = g(X, Y)$, given that $X, Y \sim N(0,1)$ and are independent.

Let's find $(x, y) = g^{-1}(r, \theta)$:

$x = g_1^{-1}(r, \theta) =$

$y = g_2^{-1}(r, \theta) =$

$$H(r, \theta) = \frac{\partial g^{-1}}{\partial (r, \theta)} =$$

$$\left| \det(H(r, \theta)) \right| =$$

$f_{X,Y}(x, y) =$

$f_{R,\Theta}(r,\theta) =$

$f_R(r) =$

R is said to have a standard Raylegh Distribution

$f_\Theta(\theta) =$

Θ has a distribution

Are R, Θ independent?

Question: What is the distribution of $Z = g(R) = R^2$, when R is a Rayleigh distribution.

In other words, what is the distribution of $R^2 = X^2 + Y^2$ if X and Y are independent standard normal r.v.'s.

Exercise : Show that if

$U \sim U(0,1)$, then the if

$X = \frac{1}{\lambda} \ln 1/U$, then $X \sim Exp(\lambda)$

Exercise: In general, if X is a r.v. with strictly increasing cdf F_X, and $U \sim U(0,1)$, then $F_X^{-1}(U)$ has the same distribution as X

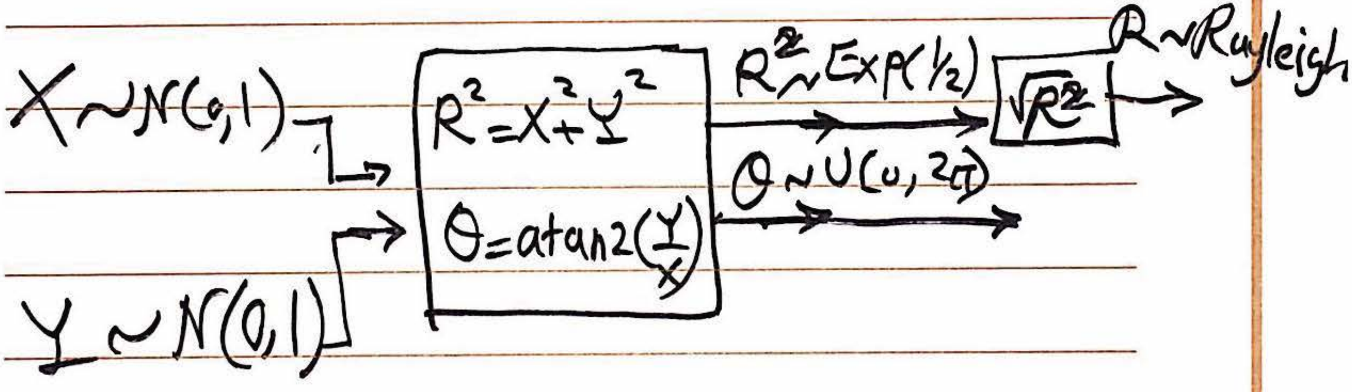
Simulation of Normal Random Variables

The previous discussion on derived distributions is very useful in simulating r.v.'s, i.e. generating samples from different distributions.

The previous exercise shows that any continuous r.v. with a strictly increasing cdf can be simulated using a uniform r.v., which can easily be generated by a simple random number generator. The CDF of a standard normal

r.v. is strictly increasing, but doesn't have ~~an~~ a closed form. Usually, the so-called Box-Muller Method is used to generate independent standard normal random variables.

The following diagram summarizes our discussion about the relationship between Rayleigh and Normal Distributions.

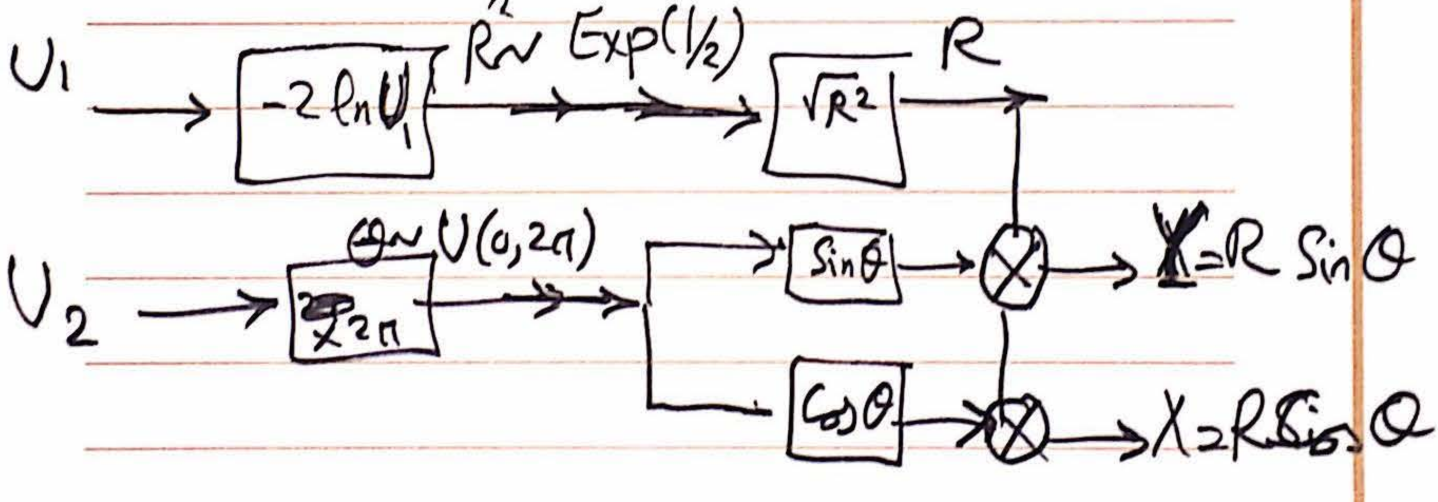


Also we learned that

$$U(0,1) \longrightarrow \underbrace{\boxed{F_X(U)}}_{-2\ln(U)} \longrightarrow X \sim Exp(1/2)$$

The Box-Muller Method does the above process in reverse:

Assume that U_1, U_2 are independent $U(0,1)$ r.v.'s



X, Y are independent $N(0,1)$ r.v.'s. (so, they are jointly normal)

Exercise: Write down the formulas for X and Y from the Box-Muller Method

Exercise: Let X and Y be independent and $X \sim$ Standard Rayleigh and $Y \sim N(0,1)$. Find the joint density of $U = \sqrt{X^2+Y^2}$ and $V = \lambda Y/X$, where $\lambda > 0, \lambda \in \mathbb{R}$.

Are U, V independent?

Rejection Sampling

(Accept-Reject Algorithm)

This algorithm doesn't need F_X^{-1}, but needs a continuous F_X.

Suppose that it is easy to sample from the pdf of Y, $g_Y(y)$.

g_Y is called the <u>instrumental pdf</u>.

Assume that

$\exists M > 1$ s.t. $\forall x$ $f_X(x) \le M g_Y(x)$

Rejection Sampling Algorithm

1. Generate Y and $U \sim U[0,1]$

2. If $U \leq \dfrac{f_X(Y)}{M g_Y(Y)}$

Then take $X = Y$ (Accept)

Else Go to Step 1. (Reject)

Theorem:

The distribution of X is f_X

Proof:

$$F_X(x) = P(X \leq x) = P\left(Y \leq x \,\middle|\, U \leq \frac{f_X(Y)}{M g_Y(Y)}\right)$$

$=$

Remark: The above result is a general procedure for "simulating" a broad class of random variables, i.e. for drawing samples/realizations from a random variable.

To draw (i.i.d) samples from X with strictly increasing F_X, one has to generate (i.i.d) $U[0,1]$ sample $U_1, U_2, \dots, U_n$. (easy)

Then $X_1 = F_X^{-1}(U_1), \dots, X_n = F_X^{-1}(U_n)$.

(U_i can be created using a random number generator)