



Mathematical Foundations for Data Science (Probability)

Transformations of Random Variables, Jacobian Formula for Transformations, Primer on Riemann Integration, Expectations

Karthik P. N.

Assistant Professor, Department of AI

Email: pnkarthik@ai.iith.ac.in

14 September 2024

Transformations of Random Variables

Maximum of Random Variables

Fix $(\Omega, \mathcal{F}, \mathbb{P})$.

Let X_1, \dots, X_n be random variables defined with respect to \mathcal{F} , with joint CDF F_{X_1, \dots, X_n} .

Let $Y_n = \max\{X_1, \dots, X_n\}$.

- Derive the CDF of Y_n .
- Simplify the CDF of Y_n when X_1, \dots, X_n are i.i.d..

Minimum of Random Variables

Fix $(\Omega, \mathcal{F}, \mathbb{P})$.

Let X_1, \dots, X_n be random variables defined with respect to \mathcal{F} , with joint CDF F_{X_1, \dots, X_n} .

- Show that $Z_n = \min\{X_1, \dots, X_n\}$ is a random variable with respect to \mathcal{F} .
- Derive the CDF of Z_n .
- Simplify the CDF of Z_n when X_1, \dots, X_n are i.i.d..

Minimum of i.i.d. Exponential Random Variables

Fix $(\Omega, \mathcal{F}, \mathbb{P})$.

Let X_1, \dots, X_n be independent, with $X_i \sim \text{Exponential}(\lambda_i)$ for each $i \in \{1, \dots, n\}$.

Find the distribution of $Z = \min\{X_1, \dots, X_n\}$.

Sums of Random Variables

Fix $(\Omega, \mathcal{F}, \mathbb{P})$.

Let X and Y be random variables with respect to \mathcal{F} .

- In the cases when X and Y are jointly discrete/continuous, derive the PMF/PDF of $X + Y$.
- Simplify the PMF/PDF when X and Y are independent.



Sum of Two Independent Exponentials

Fix $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X \sim \text{Exponential}(\mu_1)$ and $Y \sim \text{Exponential}(\mu_2)$. Assume $X \perp\!\!\!\perp Y$.

Determine the distribution of $Z = X + Y$.

CDF and Uniform Distribution

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let X be a continuous random variable with CDF F_X .

Let $g_{F_X} : [0, 1] \rightarrow \mathbb{R}$ be defined as

$$g_{F_X}(u) = \min\{x \in \mathbb{R} : F_X(x) \geq u\}, \quad u \in [0, 1].$$

- Show that $U = F_X(X)$ is uniformly distributed on $[0, 1]$.
- If $U \sim \text{Unif}([0, 1])$, show that $g_{F_X}(U)$ has the same distribution as X .

This is the basis for generating a random sample from a custom CDF F_X on a computer.

General Transformations and the Jacobian Formula

General Transformations

Fix $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be a continuous random variable with PDF f_X .

Given $g : \mathbb{R} \rightarrow \mathbb{R}$ that is **monotone** and **differentiable** with non-zero derivative throughout its domain, what is the PDF of $Y = g(X)$?

Note that g admits an inverse, say g^{-1} .

General Transformations: g Monotone Increasing

$$\begin{aligned}F_Y(y) &= \mathbb{P}(\{Y \leq y\}) \\&= \mathbb{P}(\{g(X) \leq y\}) \\&= \mathbb{P}(\{X \leq g^{-1}(y)\}) \\&= \int_{-\infty}^{g^{-1}(y)} f_X(x) dx \\&= \int_{-\infty}^y \frac{f_X(g^{-1}(v))}{g'(g^{-1}(v))} dv\end{aligned}$$

make substitution $g(x) = v$

It thus follows that

$$f_Y(y) = \begin{cases} \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}, & y \in \text{Range}(g), \\ 0, & y \notin \text{Range}(g). \end{cases}$$

General Transformations: g Monotone Decreasing

$$\begin{aligned}F_Y(y) &= \mathbb{P}(\{Y \leq y\}) \\&= \mathbb{P}(\{g(X) \leq y\}) \\&= \mathbb{P}(\{X \geq g^{-1}(y)\}) \\&= \int_{g^{-1}(y)}^{+\infty} f_X(x) dx \\&= \int_y^{+\infty} \frac{f_X(g^{-1}(v))}{g'(g^{-1}(v))} dv\end{aligned}$$

make substitution $g(x) = v$

It thus follows that

$$f_Y(y) = \begin{cases} \frac{f_X(g^{-1}(y))}{-g'(g^{-1}(y))}, & y \in \text{Range}(g), \\ 0, & y \notin \text{Range}(g). \end{cases}$$

General Transformations: g Monotone, Differentiable

When g is monotone and differentiable throughout its domain:

$$f_Y(y) = \begin{cases} \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}, & y \in \text{Range}(g), \\ 0, & y \notin \text{Range}(g). \end{cases}$$

Examples

- Let $X \sim \mathcal{N}(0, 1)$.
Derive the PDF of $Y = e^X$ from first principles and using the transformation formula.

General Transformations: g Piecewise Monotone, Differentiable

When g is piecewise monotone and differentiable

Suppose that I_1, \dots, I_n is a partition of \mathbb{R} , and $g : \mathbb{R} \rightarrow \mathbb{R}$ is piecewise monotone and differentiable with non-zero derivative on I_i for each $i \in \{1, \dots, n\}$. Let h_i denote the inverse of g on I_i . If X is a continuous random variable with PDF f_X , then the PDF of $Y = g(X)$ is given by

$$f_Y(y) = \sum_{i=1}^n \frac{f_X(h_i(y))}{\left| g'(h_i(y)) \right|} \mathbf{1}_{g(I_i)}(y).$$



Example

Suppose $X \sim \mathcal{N}(0, 1)$.

Derive the PDF of $Y = X^2$ from first principles and using the transformation formula.

General Transformations: Multivariate Case

Fix $(\Omega, \mathcal{F}, \mathbb{P})$.

Let X_1, \dots, X_n be **jointly continuous** random variables with joint PDF f_{X_1, \dots, X_n} .

Let $Y_i = g_i(X_1, \dots, X_n)$ for $i \in \{1, \dots, n\}$, where g_1, \dots, g_n are smooth¹ functions.

Derive the joint PDF of Y_1, \dots, Y_n .

¹We will assume that g_1, \dots, g_n are differentiable with continuous first-order partial derivatives.

Jacobian Matrix and Jacobian

Definition (Jacobian Matrix and Jacobian)

Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be defined by

$$g(x_1, \dots, x_n) = \begin{pmatrix} g_1(x_1, \dots, x_n) \\ \vdots \\ g_n(x_1, \dots, x_n) \end{pmatrix}$$

for some smooth functions g_1, \dots, g_n .

The **Jacobian matrix** of the mapping g at the point (x_1, \dots, x_n) is defined as

$$J_g(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1} & \dots & \frac{\partial g_n}{\partial x_n} \end{pmatrix}$$

The **Jacobian** of g at any point (x_1, \dots, x_n) is simply equal to $\det(J_g(x_1, \dots, x_n))$.

Jacobi's Transformation Formula

Fix $(\Omega, \mathcal{F}, \mathbb{P})$.

Let X_1, \dots, X_n be jointly continuous with joint PDF f_{X_1, \dots, X_n} .

Jacobi's Transformation Formula

Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be one-one, differentiable with continuous first-order partial derivatives, and non-zero Jacobian throughout its domain. Let the individual components of g be denoted by g_1, \dots, g_n . Let $Y_i = g_i(X_1, \dots, X_n)$ for all $i \in \mathbb{N}$. Then, the joint PDF of $Y = (Y_1, \dots, Y_n)$ is given by

$$f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = \begin{cases} \frac{f_{X_1, \dots, X_n}(g^{-1}(y_1, \dots, y_n))}{|J_g(g^{-1}(y_1, \dots, y_n))|}, & (y_1, \dots, y_n) \in \text{Range}(g), \\ 0, & (y_1, \dots, y_n) \notin \text{Range}(g). \end{cases}$$

Remark: For $g : \mathbb{R} \rightarrow \mathbb{R}$,

$$J_g(x) = g'(x), \quad x \in \mathbb{R}.$$

Example

Let X and Y be independent exponential random variables with parameter λ .
Derive the joint PDF of $Y_1 = X_1$ and $Y_2 = X_1 + X_2$.
Also deduce the conditional PDF of Y_1 , “conditioned on” the event $\{Y_2 = y\}$.

Jacobi's Transformation Formula: g Piecewise Differentiable with Non-Zero Jacobian

When g is piecewise differentiable with non-zero Jacobian

Suppose that I_1, \dots, I_n is a partition of \mathbb{R}^n , and $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is one-one, differentiable with continuous first-order partial derivatives, and has non-zero Jacobian on I_i for each $i \in \{1, \dots, n\}$. Let h_i denote the inverse of g on I_i . Let X_1, \dots, X_n be jointly continuous with joint PDF f_{X_1, \dots, X_n} . Let the individual components of g be g_1, \dots, g_n . Let $Y_i = g_i(X_1, \dots, X_n)$ for each $i \in \mathbb{N}$. Then the joint PDF of $Y = (Y_1, \dots, Y_n)$ is given by

$$f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = \sum_{i=1}^n \frac{f_{X_1, \dots, X_n}(h_i(y_1, \dots, y_n))}{\left| J_g(h_i(y_1, \dots, y_n)) \right|} \mathbf{1}_{g(I_i)}(y_1, \dots, y_n).$$

Integration and Expectations



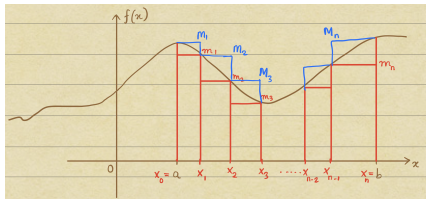
A Primer on Riemann Integration

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given.

Let $a, b \in \mathbb{R}$, $a < b$, be given.

What is the mathematical interpretation of $\int_a^b f(t) dt$?

A Primer on Riemann Integration



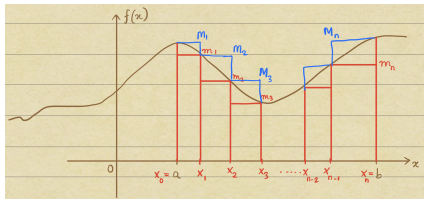
- Create a **partition** $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$, where $x_0 = a$ and $x_n = b$
- Define the **lower Riemann sum** as the total area under the red rectangles, i.e.,

$$L(f, \mathcal{P}_n) = \sum_{i=1}^n m_i \cdot |x_i - x_{i-1}|.$$

- Define the **upper Riemann sum** as the total area under the blue rectangles, i.e.,

$$U(f, \mathcal{P}_n) = \sum_{i=1}^n M_i \cdot |x_i - x_{i-1}|.$$

A Primer on Riemann Integration



- It is easy to see that for all n ,

$$L(f, \mathcal{P}_n) \leq L(f, \mathcal{P}_{n+1}), \quad U(f, \mathcal{P}_n) \geq U(f, \mathcal{P}_{n+1})$$

- If $\lim_{n \rightarrow \infty} L(f, \mathcal{P}_n) = \lim_{n \rightarrow \infty} U(f, \mathcal{P}_n)$, then this common limit is denoted $\int_a^b f(t) dt$

What if $\lim_{n \rightarrow \infty} L(f, \mathcal{P}_n) \neq \lim_{n \rightarrow \infty} U(f, \mathcal{P}_n)$?

Consider the function

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

For the above function:

- $L(f, \mathcal{P}_n) = 0$ for all n
- $U(f, \mathcal{P}_n) = 1$ for all n

Remedy:

A more general theory of integration, proposed by Lebesgue!

Lebesgue's Integration Theory

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be a random variable with respect to \mathcal{F} .

Objective

To build the necessary machinery on Lebesgue's theory of integration, so as to be able to interpret an abstract integral of the form

$$\int_A X d\mathbb{P}, \quad A \in \mathcal{F}.$$

Programme:

- Definition of the abstract integral for “simple” random variables
- Definition of the abstract integral for non-negative random variables
- Definition of the abstract integral for arbitrary random variables

Simple Random Variables

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be a random variable defined with respect to \mathcal{F} .

Definition: Simple Random Variable

A random variable X is called a **simple** random variable if it can be expressed in the form

$$X(\omega) = \sum_{i=1}^n a_i \mathbf{1}_{A_i}(\omega), \quad \omega \in \Omega,$$

for some $a_1, \dots, a_n \geq 0$ and $A_1, \dots, A_n \in \mathcal{F}$.

Example: Consider $X : \mathbb{R} \rightarrow \mathbb{R}$ defined as

$$X(\omega) = \mathbf{1}_{[0,1]}(\omega) + \frac{3}{2} \mathbf{1}_{[1,3]}(\omega), \quad \omega \in \Omega.$$

Here, X can also be represented as $X(\omega) = \mathbf{1}_{[0,3]}(\omega) + \frac{1}{2} \mathbf{1}_{[1,3]}(\omega)$.

Canonical Representation of a Simple Random Variable

Definition (Canonical Representation of a Simple Random Variable)

A simple random variable X is said to be in **canonical representation** if

$$X(\omega) = \sum_{i=1}^n a_i \mathbf{1}_{A_i}(\omega), \quad \omega \in \Omega,$$

where $a_1, \dots, a_n \geq 0$ are **distinct**, and $A_1, \dots, A_n \in \mathcal{F}$ are **disjoint**.

Integral of a Simple Random Variable

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be a random variable with respect to \mathcal{F} .

For a simple random variable X in its canonical form

$$X(\omega) = \sum_{i=1}^n a_i \mathbf{1}_{A_i}(\omega), \quad \omega \in \Omega,$$

we define $\int_{\Omega} X d\mathbb{P}$ as

$$\int_{\Omega} X d\mathbb{P} := \sum_{i=1}^n a_i \mathbb{P}(A_i).$$

The quantity $\int_{\Omega} X d\mathbb{P}$ is called the **expectation** of X under the probability measure \mathbb{P} .
Expectation of X is more commonly denoted as $\mathbb{E}[X]$.

Non-Negative Random Variables

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be any random variable with respect to \mathcal{F} such that

$$X(\omega) \geq 0 \quad \forall \omega \in \Omega.$$

Let

$$\mathcal{S}(X) := \left\{ q : \Omega \rightarrow \mathbb{R} : q \text{ simple}, q(\omega) \leq X(\omega) \quad \forall \omega \in \Omega \right\}.$$

Then, the **expectation** of the non-negative random variable X under \mathbb{P} is defined as

$$\mathbb{E}[X] = \int_{\Omega} X d\mathbb{P} := \max_{q \in \mathcal{S}(X)} \int_{\Omega} q d\mathbb{P}.$$

Remark: It is possible that $\mathbb{E}[X] = +\infty$.

Arbitrary Random Variables

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be any random variable with respect to \mathcal{F} .

Define

$$X_+(\omega) := \max\{X(\omega), 0\}, \quad \omega \in \Omega, \quad \quad X_-(\omega) := -\min\{X(\omega), 0\}, \quad \omega \in \Omega.$$

Clearly, both X_+ and X_- are non-negative random variables with respect to \mathcal{F} .

We define the expectation of X under \mathbb{P} as

$$\mathbb{E}[X] = \int_{\Omega} X d\mathbb{P} := \mathbb{E}[X_+] - \mathbb{E}[X_-],$$

provided $\min\{\mathbb{E}[X_+], \mathbb{E}[X_-]\} < +\infty$.

Arbitrary Random Variables

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be any random variable with respect to \mathcal{F} .

For any event $A \in \mathcal{F}$, we define the abstract integral $\int_A X d\mathbb{P}$ as

$$\int_A X d\mathbb{P} = \int_{\Omega} (X \cdot \mathbf{1}_A) d\mathbb{P},$$

provided the right-hand side is well-defined (i.e., not of the form $\infty - \infty$).

Properties of Expectations

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

1. If $A \in \mathcal{F}$ such that $\mathbb{P}(A) = 0$, then for any random variable X , we have

$$\mathbb{E}[X \cdot \mathbf{1}_A] = 0.$$

2. $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.
3. $\mathbb{E}[aX] = a \mathbb{E}[X]$ for all $a \in \mathbb{R}$.

Expectations for Discrete and Continuous Random Variables

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let $X : \Omega \rightarrow \mathbb{R}$ be a random variable with respect to \mathcal{F} .

- If X is a discrete random variable with PMF p_X , then

$$\mathbb{E}[X] = \sum_x x \cdot p_X(x).$$

- If X is a continuous random variable with PDF f_X , then

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x \cdot f_X(x) dx.$$