Econ 2110, fall 2016, Part Ib
Review of Probability Theory

Maximilian Kasy

Department of Economics, Harvard University

Roadmap

- Ia
  - Basic definitions
  - Conditional probability and independence
- Ib
  - Random Variables
  - Expectations
  - Transformation of variables
- Ic
  - Selected probability distributions
  - Inequalities

Part Ib

Random Variables

- Let $(\Omega, \mathscr{F}, P)$ be a probability space.
- A Random Variable X is a function
  that maps outcomes $\omega \in \Omega$ to real numbers

$$X : \Omega \mapsto \mathbb{R}.$$

- Let $A_X \subset \mathbb{R}$.
- Then the probability of the event that $X \in A_X$ is

$$\begin{aligned} & P(X \in A_X) \\ = & P_X(A_X) \\ = & P(\{\omega : X(\omega) \in A_X\}) \end{aligned}$$

- The probability measure P *induces* the probability measure P_X on the real line.
- equivalent notation:

$$
\begin{aligned}
&P(X \geq 3) \\
=&P_X([3;\infty)) \\
=&P(\{\omega : X(\omega) \geq 3\})
\end{aligned}
$$

Example

- Interviewing a random resident of Boston
- $\Omega =$ set of all residents
- $\mathscr{F} =$ is set of all subsets of Ω
- $P(\{\omega\}) = 1/N$ for all residents ω,
  where $N = |\Omega|$, the population size of Boston
- $X(\omega) =$ age of person ω
- $Y(\omega) =$ her income

Measurability

- Technical issue: we must ensure that $\{\omega : X(\omega) \in A_X\} \in \mathscr{F}$ for all A_X under consideration.
- consider $\mathscr{F}_X$ is the smallest σ-algebra that contains all intervals of the form $(-\infty, a)$ for $a \in \mathbb{R}$
- require that for all $A_X \in \mathscr{F}_X$, $\{\omega : X(\omega) \in A_X\} \in \mathscr{F}$.
  "*X* is a measurable function"
- under these conditions $(\mathbb{R}, \mathscr{F}_X, P_X)$ is a probability space.

Example

- Recall the health insurance example
- $\Omega = \{YH, YS, OH, OS\}$
- insurance premium X has to condition on public information
- $\Rightarrow X$ has to be measurable with respect to

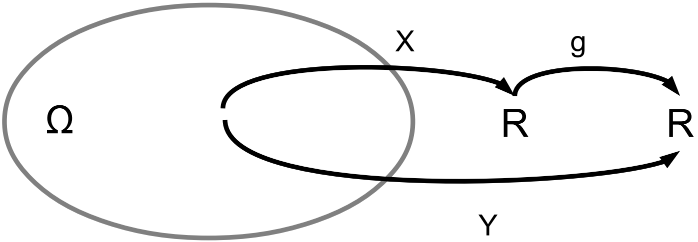
$$\mathscr{F} = \{\varnothing, \{YH, YS\}, \{OH, OS\}, \Omega\}$$

Functions of a Random Variable

- Let $g : \mathbb{R} \mapsto \mathbb{R}$ be a function, and X a random variable.
- Then
$$Y = g(X)$$
is a random variable.
- Example:
$X =$ years of education
$Y =$ attended some college = $\mathbf{1}(X > 12)$
- Probability of the event $Y \in A_Y \in \mathscr{F}_Y = \mathscr{F}_X$:

$$\begin{aligned} P(Y \in A_Y) &= P_Y(A_Y) \\ &= P_X(\{x \in \mathbb{R} : g(x) \in A_Y\}) \\ &= P(\{\omega : g(X(\omega)) \in A_Y\}) \end{aligned}$$

Figure: Functions of random variables



Distribution functions (scalar case)

- The **cumulative distribution function** (CDF) of a random variable X is defined as $F_X : \mathbb{R} \mapsto \mathbb{R}$

$$
\begin{aligned}
F_X(x) :=& P(X \leq x) \\
=& P_X((-\infty, x]) \\
=& P(\{\omega : X(\omega) \leq x\})
\end{aligned}
$$

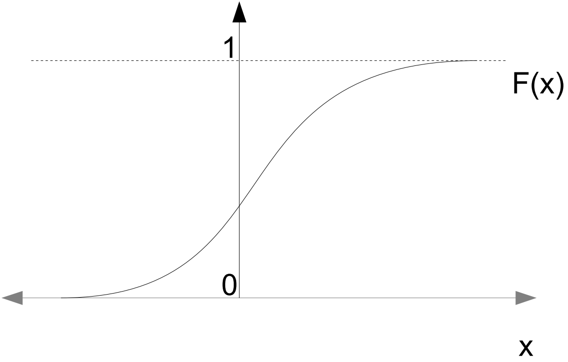
- Example:
  $X =$ income
  $F_X(20.000) =$ share of population with incomes below 20k

Practice problem

Show that , for $x_2 \geq x_1$,

$$F_X(x_2) - F_X(x_1) = P(x_1 < X \leq x_2).$$

Figure: cumulative distribution function



Properties of the CDF

1. $F_X(\cdot)$ is non-decreasing
2. $$\lim_{x \to -\infty} F_X(x) = 0$$
   $$\lim_{x \to \infty} F_X(x) = 1$$
3. $F_X(\cdot)$ is right-continuous everywhere:
   For all $x \in \mathbb{R}$,
   $$\lim_{h \downarrow 0} F_X(x+h) = F_X(x).$$

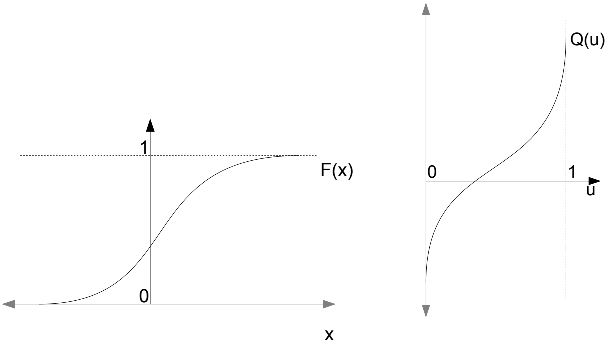
Practice problem

Show that properties 1 and 2 hold, based on the definition of a CDF and the properties of probability measures.

Quantiles

- quantiles are given by the inverse of the cumulative distribution function
$$Q(u) := \inf\{x : F(x) \geq u\}$$
- if F is invertible:
$$Q(u) = F^{-1}(u)$$
- what's the value of x, such that a population share of u is below that x?
- Example: what's the income such that 99% of the population are below that income?
- special case: **median**, $Q(0.5)$

Figure: CDF and quantile function



- For any function F that satisfies the three properties we showed, one can construct a random variable whose distribution is F:
- Let U be uniformly distributed on $[0,1]$ ($F_U(u) = u$),
- Let

$$Y = Q(U),$$

  where Q is the quantile function corresponding to F
- Then

$$F_Y(y) = P(F^{-1}(U) \leq y) = P(U \leq F(y)) = F(y)$$

- Useful for **simulations**!
- The CDF uniquely determines the probability measure P_X

Discrete random variables

- If F_X is constant except for a countable number of points $x_1, x_2, \ldots$ (i.e., F is a step function)
- then X is a discrete random variable.
- the size of the jump

$$p_i = F_X(x_i) - \lim_{h \downarrow 0} F_X(x_i - h)$$

is the probability that X takes on the value x_i:

$$P(X = x_i) = p_i.$$

- Let
$$f_X(x) = p_i$$
if $x = x_i$ and 0 otherwise.
- Then f_X is the **probability mass function** (pmf) of X
- we get
$$P(x_1 < X \leq x_2) = \sum_{x_1 < x \leq x_2} f_X(x).$$
- Examples:
  - Coinflip:
  $$f_X(0) = f_X(1) = \frac{1}{2}$$
  - Years of education (completed):
  $f_X(y) =$ share of population with exactly y years of edu

Continuous random variables

- If the CDF can be written as

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

  for some function $f_X(x)$
- then X is called a continuous random variable
- At continuity points of f_X, it then must be that

$$f_X(x) = dF_X(x)/dx$$

  by the Fundamental Theorem of Calculus
- The function f_X is called the **probability density function** of X

- for $x_2 \geq x_1$

$$\begin{aligned} P(x_1 < X \leq x_2) &= F_X(x_2) - F_X(x_1) \\ &= \int_{x_1}^{x_2} f_X(x)dx \end{aligned}$$

- Also, $P(X = x) = \int_x^x f_X(u)du = 0$ for a continuous RV.
- Examples: (approximately continuous)
  - Income
  - Hourly wage
  - Prices
  - Quantities
  - Time spent unemployed

Example

- Suppose
$$F_X(x) = 1 - e^{-x}$$
for $x \geq 0$ and $F_X(x) = 0$ for $x < 0$
- This is called the **exponential** distribution.
- Probability density of X:
$$f_X(x) = \frac{dF_X(x)}{dx} = \begin{cases} \frac{d(1-e^{-x})}{dx} = e^{-x} & \text{for } x \geq 0 \\ \frac{d0}{dx} = 0 & \text{for } x < 0 \end{cases}$$
- The **support** (set of points with positive pdf) of X is $[0, \infty)$.

Bivariate Distribution Functions

- 2 dimensional vector of random variables (X, Y) is a (measurable) mapping from Ω to $\mathbb{R}^2$.
- joint CDF

$$\begin{aligned} F_{X,Y}(x,y) &= P(X \leq x, Y \leq y) \\ &= P(\{\omega : X(\omega) \leq x\} \cap \{\omega : Y(\omega) \leq y\}) \\ &= P_{X,Y}((-\infty, x] \times (-\infty, y]) \end{aligned}$$

- Example: $(X, Y) =$ age and income
  $F(40, 30.000) =$ share of population with
  age ≤ 40 *and* income $\leq$ 30k

- (X, Y) is a discrete random vector if

$$F_{X,Y}(x,y) = \sum_{u \leq x} \sum_{v \leq y} f_{X,Y}(u,v),$$

  where $f_{X,Y}(x,y) = P(X = x, Y = y)$.
- (X, Y) is a continuous random vector if

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v) dv du$$

  for some function $f_{X,Y} : \mathbb{R}^2 \mapsto \mathbb{R}$.
- As in the scalar case, at continuity points of $f_{X,Y}$, $f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$.

Marginal Distribution

- Suppose we are given $F_{X,Y}(x,y)$ and want to recover $F_X(x)$. Then

$$\begin{aligned} F_X(x) &= P_{X,Y}((-\infty,x]\times(-\infty,\infty)) \\ &= \lim_{y\to\infty} F_{X,Y}(x,y) \end{aligned}$$

- Intuition: $P(X\leq x) = P(X\leq X \text{ and } Y\leq\infty)$

- Also

$$
\begin{aligned}
f_X(x) &= \sum_y f_{X,Y}(x,y) \quad \text{in the discrete case} \\
f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \quad \text{in the continuous case}
\end{aligned}
$$

Practice problem

Prove this for the discrete case.

- F_X and $f_X(x)$ are called 'marginal distribution' and 'marginal density'

Independence of random variables

- X and Y are independent if and only if

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \quad \text{for all} \quad x, y$$

- This implies

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

  for all x and y.
  (in both the discrete and continuous case)
- X and Y can only be independent if the support of X does not depend on Y and vice versa

- X and Y are independent if (and only if) $f_{X,Y}(x,y)$ can be written as a product of two nonnegative functions g_x and g_y

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

- where $g_x(x)$ does not depend on y and $g_y(y)$ does not depend on x
- if X and Y are independent, then so are $h(X)$ and $g(Y)$, for any choice of (measurable) functions h and g.

Conditional distributions

- Let X and Y be discrete.
  Let x be such that $f_X(x) > 0$.
- Then

$$f_{Y|X}(y|x) := P(Y = y|X = x) = \frac{f_{X,Y}(x,y)}{f_X(x)}.$$

- $f_{Y|X}(y|x)$ is called the **conditional pdf** of Y given $X = x$.
- Properties:

$$\begin{aligned} f_{Y|X}(y|x) &\geq 0 \\ \sum_y f_{Y|X}(y|x) &= \sum_y \frac{f_{X,Y}(x,y)}{f_X(x)} \\ &= \frac{f_X(x)}{f_X(x)} = 1 \end{aligned}$$

- $\Rightarrow f_{Y|X}(y|x)$ is a well defined pdf of a discrete RV.

- continuous random variables,
  for any x such that $f_X(x) > 0$:

$$f_{Y|X}(y|x) := \frac{f_{X,Y}(x,y)}{f_X(x)}$$

- **conditional pdf** of Y given $X = x$
- as long as $f_X(x) > 0$, $f_{Y|X}(y|x)$ is a well defined pdf:

$$f_{Y|X}(y|x) \geq 0$$
$$\int_{-\infty}^{\infty} f_{Y|X}(y|x)dy = 1$$

- The **conditional cdf** is

$$F_{Y|X}(y|x) = P(Y \leq y|X = x)$$
$$= \int_{-\infty}^{y} f_{Y|X}(v|x)dv$$
$$\text{or} \quad \sum_{v \leq y} f_{Y|X}(v|x)$$

- Note:

$$F_{Y|X}(y|x) \neq F_{Y,X}(y,x)/F_X(x)!$$

- For **independent** random variables, $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, so that

$$f_{Y|X}(y|x) = f_Y(y).$$

Expectations

- expectation of a discrete random variable:

$$E[X] = \sum_x x f_X(x)$$

  if $\sum_x |x| f_X(x) < \infty$.
  Otherwise, the expectation is said not to exist.
- expectation of a continuous random variable:

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

  if $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$.
  Otherwise, the expectation is said not to exist.
- Riemann-Stieltjes integral – this can be summarized as

$$E[X] = \int_{-\infty}^{\infty} x dF_X(x).$$

Linearity

- sums and integrals are linear, so is the expectation:
- For a random variable X and real numbers a and b

$$E[aX + b] = aE[X] + b$$

- provided the expectations of X exists.

Expectation of $g(X)$

- two ways to get $E[Y]$ for $Y = g(X)$
- $$\begin{aligned} E[Y] &= \int_{-\infty}^{\infty} y dF_Y(y) \\ E[g(X)] &= \int_{-\infty}^{\infty} g(x) dF_X(x) \end{aligned}$$
- proof for the discrete case:

$$\begin{aligned} \sum_y y f_Y(y) &= \sum_y y \left(\sum_x \mathbf{1}[g(x) = y] f_X(x)\right) \\ &= \sum_x \sum_y y \mathbf{1}[g(x) = y] f_X(x) \\ &= \sum_x g(x) f_X(x) \end{aligned}$$

Practice problem

1. Suppose X takes on the values $\{-1, 0, 1\}$ with probability 1/3. Let $Y = g(X) = X^2$.
   - What is the pdf of Y?
   - Calculate the expectation of Y in two ways.
2. Suppose X is distributed uniformly on [0,1] ($F_X(x) = x$ for $x \in [0, 1]$).
   - Calculate $E[X]$.
   - Calculate $E[X^2]$.

- **Probabilities as expectations:**

$$\begin{aligned}
&P(X \in A_X) \\
=&E[\mathbf{1}(X \in A_X)] \\
=&\int_{-\infty}^{\infty} \mathbf{1}(x \in A_X) dF_X(x)
\end{aligned}$$

- Expectation of a **function of several variables**:
  Let $Z = h(X, Y)$, (X, Y) continuous. Then

$$\begin{aligned}
E[Z] &= \int_{-\infty}^{\infty} z f_Z(z) dz \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f_{X,Y}(x, y) dx dy
\end{aligned}$$

- by **linearity**, for any two random variables X and Y and real numbers a and b,

$$E[aX + bY] = aE[X] + bE[Y].$$

- if X and Y are **independent**,

$$\begin{aligned} E[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dxdy \\ &= \int_{-\infty}^{\infty} xf_X(x)dx \int_{-\infty}^{\infty} yf_Y(y)dy \\ &= E[X]E[Y] \end{aligned}$$

Moments

- kth moment of X: $E[X^k]$
- first moment: **mean**, $\mu = E[X]$
  – a measure of location.
- kth centered moment: $E[(X-\mu)^k]$.
- second centered moment: **variance** σ^2
  – a measure of spread

$$\sigma^2 = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] = E[X^2] - \mu^2$$

- square root of the variance, σ: **standard deviation.**

- All odd centered moments are zero for RVs with symmetric distribution
  (i.e. $f_X(\mu + x) = f_X(\mu - x)$ for all x).
- third centered moment: '**skewness**'
- fourth centered moment: '**kurtosis**'

Moments for vector valued random variables

- suppose $X = X = (X_1, \cdots, X_n)'$
- mean $\mu = E[X]$, is defined as the $n \times 1$ vector

$$\mu = \begin{pmatrix} E[X_1] \\ \vdots \\ E[X_2] \end{pmatrix}$$

- **covariance** matrix:

$$\Sigma = E[(X - \mu)(X - \mu)']$$

- covariance between X_i and X_j:
  $\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$
- Note that $\sigma_{ij} = \sigma_{ji}$, so that Σ is **symmetric** ($\Sigma' = \Sigma$)

- Let α and β be $n \times 1$ non-stochastic vectors.
-

$$E[\alpha' X] = \alpha' \mu_X$$
$$Var[\alpha' X] = \alpha' \Sigma \alpha \geq 0$$

- $\Rightarrow \Sigma$ is **positive semi-definite.**
- **covariance** between $\alpha' X$ and $\beta' X$:

$$E[(\alpha' X - \alpha' \mu_X)(\beta' X - \beta' \mu_X)] = \alpha' \Sigma \beta$$

Correlation

- Let $\rho_{ij} = \sigma_{ij} / \sqrt{\sigma_{ii}\sigma_{jj}}$. Then ρ_{ij} is called the **correlation** between X_i and X_j.
- Since Σ is positive semi-definite, so is

$$V = \begin{pmatrix} \sigma_{ii} & \sigma_{ij} \\ \sigma_{ji} & \sigma_{jj} \end{pmatrix}$$

- thus

$$0 \leq |V| = \sigma_{ii}\sigma_{jj} - \sigma_{ij}^2$$

- and $-1 \leq \rho_{ij} \leq 1$.
- If X_i and X_j are independent, then $\rho_{ij} = \sigma_{ij} = 0$ (the converse is false)

Conditional Expectations

- conditional expectation of Y given $X = x$ (for $f_X(x) > 0$):
- the expectation of Y with respect to the conditional probability density $f_{Y|X}(y|x)$
- continuous case:

$$\mu_Y(x) = E[Y|X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy$$

  which is a function that depends on x.
- Viewed as such, $\mu_Y(x) = E[Y|X = x]$ is sometimes called a **regression function**.

- Examples:
    - average income Y for women ($X = 0$) and men ($X = 1$)
    - share of unemployed ($Y = 1$) for people of different ages X
- Since $\mu_Y(x) = E[Y|X = x]$ is a function $\mathbb{R} \mapsto \mathbb{R}$,
- $\mu_Y(X) = E[Y|X]$ is a random variable
  – functions of random variables are random variables.
- don't need to worry about a definition of $E[Y|X = x]$ for x with $f_X(x) = 0$, since the probability of observing X such x is zero.

Law of iterated expectations

Theorem

For Random Variables X and Y, $E[E[Y|X]] = E[Y]$
(provided the expectations exist.)

Proof for the continuous case:

$$
\begin{aligned}
E[E[Y|X]] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy f_X(x) dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{Y|X}(y|x) f_X(x) dy dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{X,Y}(x,y) dy dx \\
&= \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy \\
&= \int_{-\infty}^{\infty} y f_Y(y) dy = E[Y]
\end{aligned}
$$

Alternative definition of conditional expectations

- Can also think of conditional expectation as an orthogonal projection
- gives useful geometric intuition!
- Space of random variables (on Ω)
  with $E[X^2] < \infty$
- equipped with inner product

$$\langle X, Y \rangle := E[X \cdot Y]$$

- so-called L^2 space

- Then $E[Y|X]$ is the orthogonal projection of Y on the space of random variables which are functions of X
- $\mu(.)$ minimizes mean squared prediction error,

$$\mu(.) = \operatorname*{argmin}_{m(.)} E[(Y - m(X))^2]$$

- implications:
  1. orthogonal projections are linear
  2. law of iterated expectations $\sim$ iterated projections
  3. regression residuals are orthogonal to predictors
- can also project on other spaces,
  eg. space of linear functions of X
  $\Rightarrow$ best linear predictor

Practice problem

- Find the coefficients $\beta = (\beta_0, \beta_1)$
- of the best linear predictor
- for Y given X,

$$\beta = \underset{b}{\operatorname{argmin}} \; E[(Y - b_0 - b_1 X)^2].$$

Transformation of Variables

- Let X be a random variable with cdf F_X.
- Let $Y = h(X)$, where $h : \mathbb{R} \mapsto \mathbb{R}$ has range $R = \{y : y = h(x), x \in \mathbb{R}\}$ and is one-to-one with inverse h^{-1}.
- What is the distribution of Y?
- Discrete case: For $y \in R$,

$$f_Y(y) = P(Y = y) = P(X = h^{-1}(y)) = f_X(h^{-1}(y))$$

and $F_Y(y) = \sum_{v \leq y, v \in R} f_X(h^{-1}(v))$.

Continuous case

- first suppose that h is increasing. For $y \in R$

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(X \leq h^{-1}(y)) \\
&= F_X(h^{-1}(y))
\end{aligned}
$$

- and

$$
\begin{aligned}
f_Y(y) &= \frac{dF_Y(y)}{dy} \\
&= \frac{dF_X(h^{-1}(y))}{dy} \\
&= f_X(h^{-1}(y))\frac{dh^{-1}(y)}{dy}
\end{aligned}
$$

- With h decreasing,

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(X \geq h^{-1}(y)) \\
&= 1 - F_X(h^{-1}(y))
\end{aligned}
$$

- so that $f_Y(y) = -f_X(h^{-1}(y))\frac{dh^{-1}(y)}{dy}$.
- combining the increasing / decreasing results, this yields

$$
f_Y(y) = f_X(h^{-1}(y))\left|\frac{dh^{-1}(y)}{dy}\right|
$$

Special case

- interesting special case: $h(x) = F_X(x)$
- suppose F_X is continuous.
- Let $F_X^{-1}(y)$ be defined as the smallest x such that $F_X(x) = y$
- Then

$$\begin{aligned} F_Y(y) =& P(F_X(X) \leq y) \\ =& P(F_X^{-1}(F_X(X)) \leq F_X^{-1}(y)) \\ =& P(X \leq F_X^{-1}(y)) \\ =& F_X(F_X^{-1}(y)) = y \end{aligned}$$

- therefore $F_X(X)$ is distributed uniformly on [0;1].

Bivariate case

- Let X_1 and X_2 be 2 random variables with joint density $f_{X1,X2}(x_1, x_2)$,
- let $h(x_1, x_2) = (h_1(x_1, x_2), h_2(x_1, x_2))$ be one-to-one,
- with inverse mapping $h^{-1}(y_1, y_2) = (h_1^{-1}(y_1, y_2), h_2^{-1}(y_1, y_2))$
- denote the range of h as R.
- on R, the pdf of $Y = (Y_1, Y_2) = h(X_1, X_2)$ is given by

$$f_{Y1,Y2}(y_1, y_2) = f_X(h^{-1}(y)) \cdot |J(y)|$$

- where the Jacobian determinant J is defined as

$$J(y_1, y_2) = \det \begin{pmatrix} \frac{\partial h_1^{-1}(y_1,y_2)}{\partial y_1} & \frac{\partial h_1^{-1}(y_1,y_2)}{\partial y_2} \\ \frac{\partial h_2^{-1}(y_1,y_2)}{\partial y_1} & \frac{\partial h_2^{-1}(y_1,y_2)}{\partial y_2} \end{pmatrix}.$$

Practice problem

We are given that the joint pdf X_1 and X_2 is $f_{X1,X2}(x_1, x_2) = e^{-(x_1+x_2)}\mathbf{1}[x_1 \geq 0]\mathbf{1}[x_2 \geq 0]$. What is distribution of

$$(Y_1, Y_2) = h(X_1, X_2) = (X_1 + X_2, X_1 - X_2)?$$

Solution:

- The range R is $R = \mathbb{R} \times \mathbb{R}$
- $h^{-1}(y_1, y_2) = (\frac{y_1+y_2}{2}, \frac{y_1-y_2}{2})$.
- Hence
$$J = \det \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = -\frac{1}{2}$$
- therefore

$$\begin{aligned} f_{Y1,Y2}(y_1, y_2) &= \tfrac{1}{2} e^{-(\frac{y_1+y_2}{2} + \frac{y_1-y_2}{2})} \mathbf{1}[\frac{y_1+y_2}{2} \geq 0] \mathbf{1}[\frac{y_1-y_2}{2} \geq 0] \\ &= \tfrac{1}{2} e^{-y_1} \mathbf{1}[y_1+y_2 \geq 0] \mathbf{1}[y_1 \geq y_2] \end{aligned}$$

- What is the implied distribution for $Z = X_1 + X_2$?
-

$$
\begin{aligned}
f_Z(z) =& f_{Y1}(y_1) \\
=& \int_{-\infty}^{\infty} f_{Y1,Y2}(y_1, y_2) dy_2 \\
=& \int_{-\infty}^{\infty} \tfrac{1}{2} e^{-y_1} \mathbf{1}[y_1 + y_2 \geq 0] \mathbf{1}[y_1 \geq y_2] dy_2 \\
=& \mathbf{1}[y_1 \geq 0] \tfrac{1}{2} e^{-y_1} \int_{-y_1}^{y_1} dy_2 \\
=& \mathbf{1}[y_1 \geq 0] y_1 e^{-y_1}
\end{aligned}
$$