Random variables

Probability and distributions

Francesco Corona

Department of Computer Science
Federal University of Ceará, Fortaleza

Random variables

Probability and distributions

Random variables

If the elements of $\mathcal{C}$ are not numbers, the sample space is dull to describe

$\rightsquigarrow$ A set of rules to represent elements c of $\mathcal{C}$ by numbers

Example

Let the toss of a coin be the random experiment

Let $\mathcal{C} = \{\texttt{H}, \texttt{T}\}$ be the sample space we associate with this experiment

- H and T are for heads and tails

Let X be a function such that $X(\texttt{T}) = 0$ and $X(\texttt{H}) = 1$

- X is a real-valued function defined on the sample space $\mathcal{C}$
- From $\mathcal{C}$, to a space of real numbers $\mathcal{D} = \{0, 1\}$

We can define a random variable and its space

Random variables

Definition

A random variable and its space/range

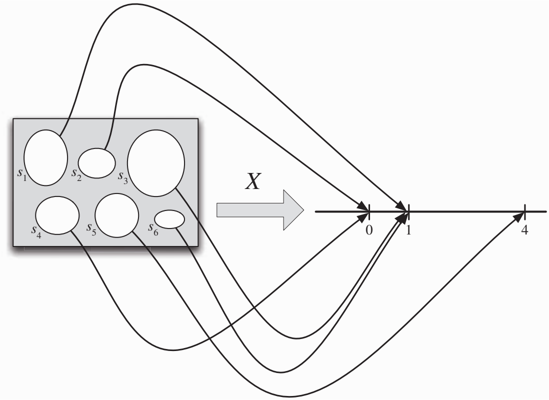
Consider a random experiment and let $\mathcal{C}$ be the sample space

A ***random variable (RV)*** *is a function X that assigns to each element $c \in \mathcal{C}$ one and only one number, $X(c) = x$*

The ***space/range*** *of X is the set of real numbers $\mathcal{D} = \{x : x = X(c), c \in \mathcal{C}\}$*

Random variables (cont.)

A random variable thus maps the sample space onto the real line



Randomness comes from choosing a random element from sample space

Random variables (cont.)

The range $\mathcal{D}$ is typically a countable set or an interval of real numbers

- RVs of the first type are said to be **discrete**
- RVs of the second time are said to be **continuous**

Example

- RV X is defined on a sample space with 6 elements ($\mathcal{C}$)
- RV X has possible values 0, 1 and 4 ($\mathcal{D}$)

Random variables (cont.)

Given a RV X, its range $\mathcal{D}$ becomes the sample space of interest

- Besides the sample space, X also induces a probability
- This probability is called the **distribution** of X

Random variables (cont.)

Let X be a discrete random variable with finite range $\mathcal{D} = \{d_1, d_2, \dots, d_m\}$

- The events of interest of the new sample space $\mathcal{D}$ are subsets of $\mathcal{D}$

We define function $p_X(d_i)$ on $\mathcal{D}$

$$p_X(d_i) = P\big[\{c : X(c) = d_i, c \in \mathcal{C}\}\big], \quad \text{for } i = 1, 2, \dots, m \tag{1}$$

Function $p_X(d_i)$ defines the **probability mass function** (**PMF**) of X

The induced probability distribution $P_X(\cdot)$ of X

$$P_X(D) = \sum_{d_i \in D} p_X(d_i), \quad D \subset \mathcal{D}$$

Random variables (cont.)

Example

Two dice roll

Let X be the sum of upfaces on a single roll of two fair and ordinary dice

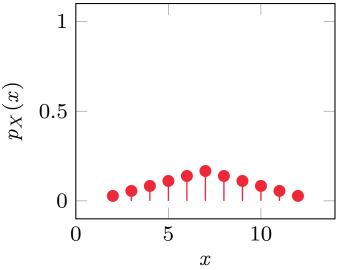
- The sample space is $\mathcal{C} = \{(i,j) : 1 \leq i,j \leq 6\}$
- As the dice are fair, $P\big[\{(i,j)\}\big] = 1/36$

- Random variable X is $X\big[(i,j)\big] = i + j$
- The range of X is $\mathcal{D} = \{2, \dots, 12\}$

Random variables (cont.)

The probability mass function PMF[1] of X is given (by enumeration) by

x	2	3	4	5	6	7	8	9	10	11	12
$p_X(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$



[1] $p_X(d_i) = P[\{c : X(c) = d_i\}]$, for $i = 1, 2, \ldots, m$.

Random variables (cont.)

The computation of probabilities regarding X follows

Suppose $B_1 = \{x : x = 7, 11\}$ and $B_2 = \{x : x = 2, 3, 12\}$

Using the values of $p_X(x)$ in the table

$$P_X(B_1) = \sum_{x \in B_1} p_X(x) = \frac{6}{36} + \frac{2}{36} = 8/36$$

$$P_X(B_2) = \sum_{x \in B_2} p_X(x) = \frac{1}{36} + \frac{2}{36} + \frac{1}{36} = 4/36$$

Random variables (cont.)

Let X be a continuous random variable, the new sample space $\mathcal{D}$ is an interval of real numbers and simple events of interest are (generally) intervals

Usually, we can determine a non-negative function $f_X(x)$ such that, for any interval $(a, b) \in \mathcal{D} \subset \mathcal{R}$, the induced probability distribution $P_X(\cdot)$ of X is

$$P_X\big[(a, b)\big] = P\big[\{c : a < X(c) < b, c \in \mathcal{C}\}\big] = \int_a^b f_X(x)\mathrm{d}x \qquad (2)$$

- This is the probability that X falls between a and b
- The area under curve $y = f_X(x)$ between a and b

Function $f_X(x)$ defines the **probability density function** of X (**PDF**)

Random variables (cont.)

Besides $f_X(x) \geq 0$, we require that

$$P_X(\mathcal{D}) = \int_{\mathcal{D}} f_X(x) \mathrm{d}x = 1$$

Total area under curve over the sample space $\mathcal{D}$ is 1

Random variables (cont.)

Example

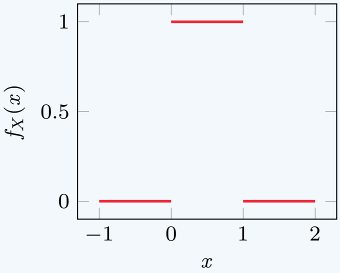
Pick a real number at random from interval $(0,1)$, let X be the number

The space of X is $\mathcal{D} = (0,1)$, the induced probability P_X is not obvious

- As the number is picked at random, it makes sense to assign

$$P_X\big[(a,b)\big] = b - a, \text{ for } 0 < a < b < 1 \tag{3}$$

Since $P_X\big[(a,b)\big] = \int_a^b f_X(x)\mathrm{d}x = b - a$, it follows that the PDF of X



$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases} \tag{4}$$

Random variables (cont.)

The probability that X is less than one-eighth or greater than seven-eighths

$$P\big[\big(X < 1/8\big) \cup \big(X > 7/8\big)\big] = \int_0^{1/8} (1)\mathrm{d}x + \int_{7/8}^{1} (1)\mathrm{d}x = 1/4$$

■

Random variables (cont.)

Remark

- $p_X(d_i) = P[\{c \in \mathcal{C} : X(c) = d_i\}]$ for $i = 1, \dots, m$
- $P_X[(a,b)] = P[\{c \in \mathcal{C} : a < X(c) < b\}] = \int_a^b f_X(x)\mathrm{d}x$

Subscript X identifies the PMF $p_X(x)$ and the PDF $f_X(x)$ with the RV

Remark

PMFs of discrete RVs and PDFs of continuous RVs are different beasts

Random variables (cont.)

Distribution functions determine the probability distribution of a RV

Definition

Cumulative distribution function

Let X be a random variable

*The **cumulative distribution function (CDF)** of RV X is defined by $F_X(x)$*

$$F_X(x) = P_X\Big[(-\infty, x]\Big] = P\Big[\{c : X(c) \leq x, c \in \mathcal{C}\}\Big] \tag{5}$$

We conveniently shorten $P\big[\{c : X(c) \leq x, c \in \mathcal{C}\}\big]$ as $P(X \leq x)$

Random variables (cont.)

Remark

$F_X(x)$ is often called the **distribution function** (**DF**) of X

- Adjective 'cumulative' indicates that probabilities are accumulated
- $F_X(x)$ adds up probabilities less than or equal to x

Random variables (cont.)

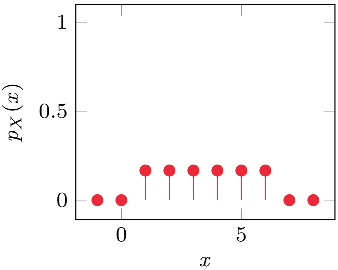
Example

Roll of a fair die

Suppose we roll an ordinary die, the dice is fair

- Let X be the observed spots upface
- The space of X is $\{1, 2, \ldots, 6\}$

The PMF of X



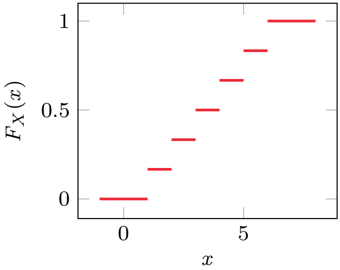
$p_X(i) = 1/6$

for $i = 1, 2, \ldots, 6$

Random variables (cont.)

- If $x < 1$, then $F_X(x) = \sum_{i \in (-\infty, 1)} p_X(i) = 0$
- If $1 \leq x < 2$, then $F_X(x) = \sum_{i \in (-\infty, 2)} p_X(i) = 1/6$
- If $2 \leq x < 3$, then $F_X(x) = \sum_{i \in (-\infty, 3)} p_X(i) = 2/6$
- ...

Continuing this way, the CDF of X is an increasing step function



The step of $F_X(x)$ is $p_X(i)$, at each i, in the range of X

⇝ Given the CDF of X, we can determine its PMF

Random variables (cont.)

Example

Pick a real number at random from interval $(0, 1)$, let X be the number

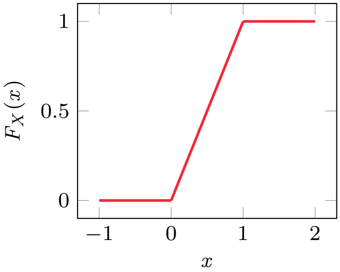
The space of X is $\mathcal{D} = (0, 1)$

What is the CDF of X?

- If $x < 0$, then $P(X \leq x) = 0$
- If $1 \leq x$, then $P(X \leq x) = 1$
- If $0 \leq x < 1$, then $P(X \leq x) = P(0 < X \leq x) = x - 0 = x$

Random variables (cont.)

Hence, the CDF if X is given by



$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \in [0, 1) \\ 1, & \text{if } x \geq 1 \end{cases} \tag{6}$$

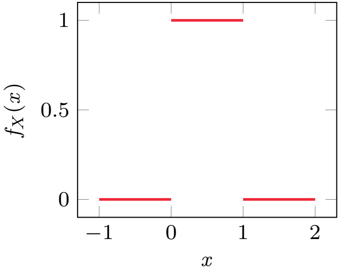
Random variables (cont.)

The connection between $F_X(x)$ and $f_X(x)$ is given by

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\mathrm{d}x, \text{ for all } x \in \mathcal{R}$$

and

$$\frac{\partial}{\partial x} F_X(x) = f_X(x), \text{ for all } x \in \mathcal{R}, \text{ except } x = 0 \text{ and } x = 1$$



$$f_X(x) = \begin{cases} 1, & x \in (0,1) \\ 0, & \text{elsewhere} \end{cases}$$

Random variables (cont.)

Let X and Y be two random variables

X and Y are equal in distribution $X \stackrel{D}{=} Y$ iff $F_X(x) = F_Y(y)$, for all $x \in \mathcal{R}$

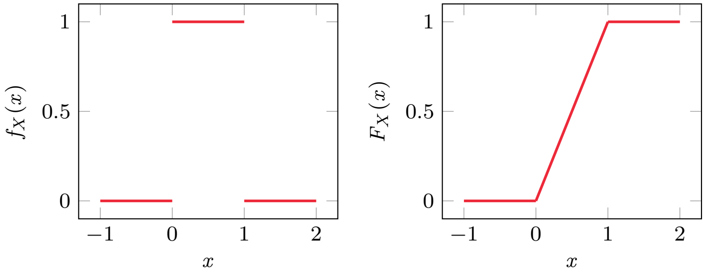
- X and Y can be equal in distribution, and yet be different otherwise

Random variables (cont.)

Example

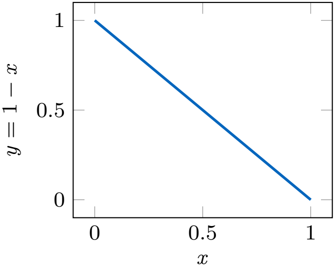
Pick a real number at random from interval $(0, 1)$, let X be the number

- The range of X is $\mathcal{D}_X = (0, 1)$



Random variables (cont.)

Define the random variable $Y = 1 - X$, clearly $Y \neq X$



The range of Y is the same as X

- Interval $\mathcal{D}_Y = (0, 1)$

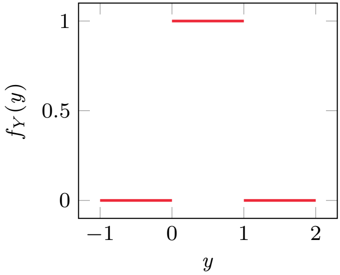
Random variables (cont.)

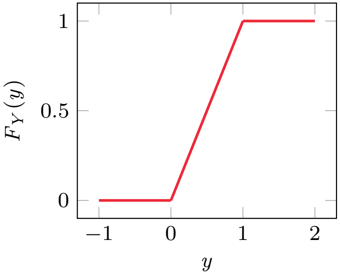
The CDF of Y is given by

$$F_Y(y) = \begin{cases} 0, & \text{if } y < 0 \\ y, & \text{if } y \in [0, 1) \\ 1, & \text{if } y \geq 1 \end{cases}$$

For $0 \leq y < 1$, we computed

$$F_Y(y) = P(Y \leq y) = P(1 - X \leq y) = P(X \geq 1 - y) = 1 - (1 - y) = y$$





$\rightsquigarrow \quad Y \stackrel{D}{=} X$

Random variables (cont.)

We have seen two CDFs that share some common features

- They are increasing functions
- The lower limit is 0, the upper limit 1

Such properties are true for CDFs in general

Theorem 1.1

Let X be a random variable

Let its cumulative distribution function be F_X

a) *For all a and b, if $a < b$, then $F(a) \leq F(b)$ (F is non-decreasing)*
b) $\lim_{x \downarrow -\infty} F(x) = 0$ *(lower-bound of F is zero)*
c) $\lim_{x \uparrow +\infty} F(x) = 1$ *(upper-bound of F is one)*
d) $\lim_{x \downarrow x_0} F(x) = F(x_0)$ *(F is right-continuous)*

Random variables (cont.)

Theorem 1.2

Let X be a random variable with F_X the CDF

Then, for $a < b$,

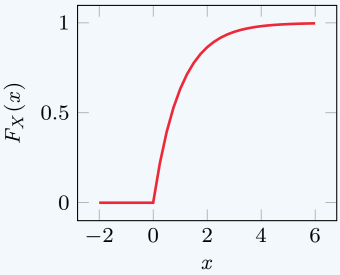
$$P(a \leq X \leq b) = F_X(b) - F_X(a)$$

Random variables (cont.)

Example

Let X be the lifetime in years of your potted flowers

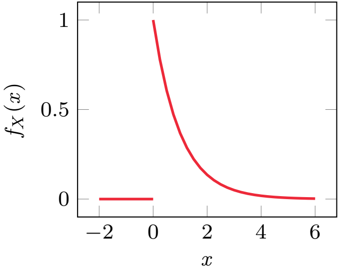
Assume that X has the CDF



$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x}, & 0 \leq x \end{cases}$$

Random variables (cont.)

The PDF of X, $\mathrm{d}[F_X(x)]/\mathrm{d}x$



$$f_X(x) = \begin{cases} e^{-x}, & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases}$$

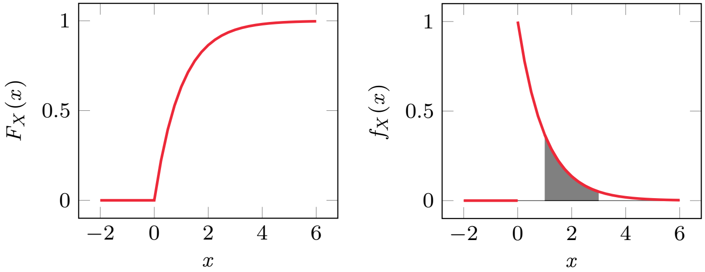
Remark

The derivative does not exist at $x = 0$

Random variables (cont.)

The probability that your flower has a lifetime between 1 and 3 years

$$P(1 < X \leq 3) = F_X(3) - F_X(1) = \int_1^3 e^{-x} \mathrm{d}x$$



The probability is found by $F_X(3) - F_X(1)$ or by evaluating the integral

- Either way, it equals $\exp(-1) - \exp(-3) = 0.318$

Random variables (cont.)

CDFs are right-continuous and monotone

CDFs possess a countable number of discontinuities

- It can be shown that the discontinuities of a CDF have mass
- If x is a point of discontinuity of F_X, then $P(X = x) > 0$

Random variables (cont.)

Theorem

For any random variable X,

$$P[X = x] = F_X(x) - F_X(x-), \quad \text{for all } x \in \mathcal{R} \tag{7}$$

We used $F_X(x-) = \lim_{z \uparrow x} F_X(z)$

Proof

For any $x \in \mathcal{R}$, we have

$$\{x\} = \bigcap_{n=1}^{\infty} \Big(x - \frac{1}{n}, x\Big],$$

$\{x\}$ is the limit of a decreasing sequence of sets

Random variables (cont.)

For a decreasing sequence of sets

$$\lim_{n \uparrow \infty} P(C_n) = P(\lim_{n \uparrow \infty} C_n) = P\Big(\bigcap_{n=1}^{\infty} C_n \Big)$$

Hence,

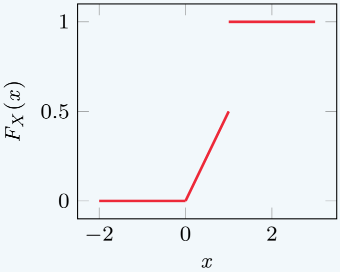
$$P(X = x) = P\Big[\bigcap_{n=1}^{\infty} \Big\{ x - \frac{1}{n} < X \leq x \Big\} \Big] = \lim_{n \uparrow \infty} P\Big[x - \frac{1}{n} < X \leq x \Big]$$
$$= \lim_{n \uparrow \infty} \Big\{ F_X(x) - F_X\left[x - (1/n)\right] \Big\} = F_X(x) - F_X(x-)$$

■

Random variables (cont.)

Example

Let the random variable X have the discontinuous CDF $F_X(x)$



$$F_X(x) = \begin{cases} 0, & x < 0 \\ x/2, & 0 \leq x < 1 \\ 1, & 1 \leq x \end{cases}$$

Then,

$$P(-1 < X \leq 1/2) = F_X(1/2) - F_X(-1) = 1/4 - 0 = 1/4$$

and

$$P(X = 1) = F_X(1) - F_X(1-) = 1 - 1/2 = 1/2$$

The value $1/2$ equals the value of the step of F_X for $x = 1$

Random variables (cont.)

The total probability associated with a random variable X of discrete type with PMF $p_X(x)$, or of the continuous type with PMF $f_X(x)$ is 1

Then, it must be true that

$$\sum_{x \in \mathcal{D}} p_X(x) = 1 \quad \text{and} \quad \int_{\mathcal{D}} f_X(x) \mathrm{d}x = 1,$$

$\mathcal{D}$ denotes the space of X

Random variables (cont.)

Remark

If we know the PMF/PDF up to a constant, then we know the PMF/PDF

Random variables (cont.)

Example

Suppose that X has the PMF

$$p_X(x|c) = \begin{cases} cx, & x = 1, 2, \ldots, 10 \\ 0, & \text{elsewhere} \end{cases},$$

for some proper constant c

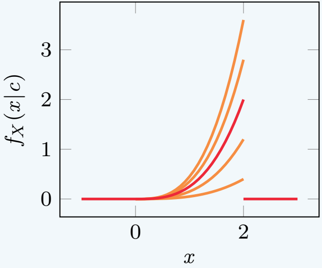
Then,

$$1 = \sum_{x=1}^{10} p_X(x) = \sum_{x=1}^{10} cx = c(1 + 2 + \cdots + 10) = 55c$$

■

Random variables (cont.)

Example



Suppose that X has the PDF

$$f_X(x|c) = \begin{cases} cx^3, & 0 < x < 2 \\ 0, & \text{elsewhere} \end{cases},$$

for some constant c

Then,

$$1 = \int_0^2 cx^3 \mathrm{d}x = c\left[\frac{x^4}{4}\right]_0^2 = 4c$$

The computation of a probability involving X follows as always

$$P\big(1/4 < X < 1\big) = \int_{1/4}^{1} \frac{1}{4}x^3 \mathrm{d}x = \frac{255}{4096} \approx 0.06$$

Discrete random variables

Random variables

Generalities

Discrete random variables

Discrete random variables

Definition

Discrete random variables

*A random variable is said to be a **discrete random variable** if its space/range is either finite or countable*

Discrete random variables

Example

Independent tosses of a coin

Consider a sequence of independent tosses of a coin

- Each flip results in a head H or a tail T
- Assume that H and T are equiprobable

The sample space $\mathcal{C}$ of the experiment consists of sequences like $\texttt{TTHHTHT}\cdots$

Let the random variable X be the number of tosses to observe the first H

- $X(\texttt{HTTHHTT}\cdots) = 1$
- $X(\texttt{THHTHHH}\cdots) = 2$
- $X(\texttt{TTHHTHT}\cdots) = 3$
- $\cdots$

The space range of X is $\mathcal{D} = \{1, 2, \dots\}$

Discrete random variables (cont.)

We have that $X = 1$ when the sequence starts with `H`

- $P(X = 1) = 1/2$

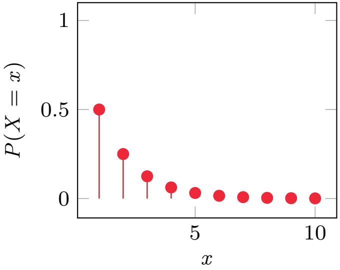
Similarly, $X = 2$ when the sequence starts with `TH`

- $P(X = 2) = (1/2)(1/2) = (1/2)^2 = 1/4$

For $X = x$, there must be a sequence of $(x - 1)$ `T`s followed by a `H`

- $P(X = x) = \underbrace{(1/2)(1/2)\cdots}_{x-1 \text{ times}}(1/2) = (1/2)^{x-1}(1/2) = (1/2)^x$

Discrete random variables (cont.)



$$P(X = x) = \big(1/2\big)^x \quad \text{for } x = 1, 2, 3, \ldots \tag{8}$$

The event 'first H is observed on a odd number of flips' has probability

$$P\big[X \in \{1, 3, 5, \cdots\}\big] = \sum_{x=1}^{\infty} \big(1/2\big)^{2x-1} = \frac{1/2}{1-(1/4)} = 2/3$$

Discrete random variables (cont.)

Probabilities regarding discrete RVs can be obtained in terms of probabilities

$$P(X = x), \text{ for } x \in \mathcal{D}$$

These probabilities determine an important probability function

Discrete random variables (cont.)

Definition

***Probability mass function** (**PMF**)*

Let X be random variable of the discrete type with range $\mathcal{D}$

*The **probability mass function** (**PMF**) of X is given by*

$$p_X(x) = P(X = x), \quad \textit{for } x \in \mathcal{D} \tag{9}$$

The PMF satisfies two properties

$$\begin{aligned} &(i) \quad 0 \leq p_X(x) \leq 1, \text{ for } x \in \mathcal{D} \\ &(ii) \quad \sum_{x \in \mathcal{D}} p_X(x) = 1 \end{aligned} \tag{10}$$

[$\rightsquigarrow$] It can be shown that the distribution of the RV is uniquely determined by a function that satisfies properties (i) and (ii) for a discrete set $\mathcal{D}$

Discrete random variables (cont.)

Let X be a discrete RV with range $\mathcal{D}$, discontinuities of $F_X(x)$ possess mass

$\rightsquigarrow$ If x is a point of discontinuity of F_X, then $P(X = x) > 0$

The range of a discrete RV and such points of positive probability are distinct

- In the range $\mathcal{D}$ of discrete random variable X, the set of points that have positive probability are said to define the **support** of RV X
- We often use $\mathcal{S}$ to indicate the support of X
- $\mathcal{S} \subset \mathcal{D}$ and it may be that $\mathcal{S} = \mathcal{D}$

Discrete random variables (cont.)

We can also obtain a relation between the PMF and CDF of a discrete RV

- If $x \in \mathcal{S}$, then $p_X(x)$ equals the size of the discontinuity of F_X at x
- If $x \notin \mathcal{S}$, then $P(X = x) = 0$ and, hence, F_X is continuous at such x

Discrete random variables

Example

A box consisting of 100 bulbs is inspected using a standardised procedure

- Five bulbs are selected at random and checked
- If they all glow, the lot is accepted

If there are 20 faulty bulbs in the lot, the probability of accepting the box

$$\binom{80}{5} \Big/ \binom{100}{5} \approx 0.32$$

Discrete random variables (cont.)

Let the RV X be the number of faulty bulbs among the inspected 5

The PMF of X is

$$p_X(x) = \begin{cases} \binom{20}{x}\binom{80}{5-x}/\binom{100}{5}, & x = 0,1,2,3,4,5 \\ 0 & \text{elsewhere} \end{cases} \tag{11}$$

The range of X, $\mathcal{D} = \{0,1,2,3,4,5\}$ is also its support $\mathcal{S}$

Transformations

Discrete random variables

Transformations

Suppose that we have a RV X of the discrete type, we know its distribution

- We are interested in another RV Y
- RV Y is some transformation of X

$\rightsquigarrow$ $Y = g(X)$

Specifically, we want to determine the distribution of Y

Transformations (cont.)

Assume X is of the discrete type with range $\mathcal{D}_X$

- The range of Y is $\mathcal{D}_Y = \{g(x) : x \in \mathcal{D}_X\}$

We consider two cases, separately

1. g is one-to-one (bijective)
2. g is not one-to-one

Transformations (cont.)

g is one-to-one

The PMF of Y

$$\begin{aligned} p_Y(y) &= P(Y = y) = P[g(X) = y] = P[X = g^{-1}(y)] \\ &= p_X[g^{-1}(y)] \end{aligned} \tag{12}$$

Transformations (cont.)

Example

Consider a sequence of independent tosses of a coin

- Each flip results in a `H` or a `T`
- `H` and `T` are equiprobable

The sample space $\mathcal{C}$ of the experiment consists of sequences like $\texttt{TTHHTHT}\cdots$

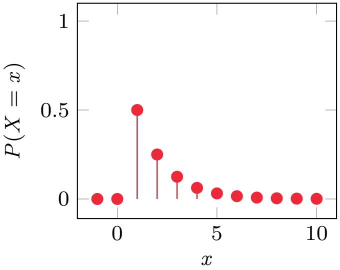
Let the random variable X be the number of tosses to observe the first `H`

- $X(\texttt{TTHHTHT}\cdots) = 3$
- $\cdots$

The range of X is $\mathcal{D}_X = \{1, 2, 3, 4, \dots\}$

Transformations (cont.)

The PMF of X was determined earlier

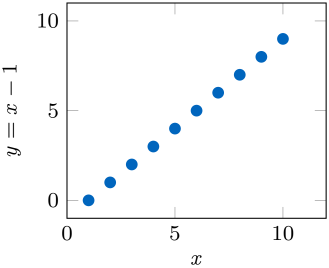


$$P(X = x) = \big(1/2\big)^x$$
$$\text{for } x = 1, 2, 3, \dots$$

Transformations (cont.)

Let $Y = X - 1$ (the number of tosses before the first H)

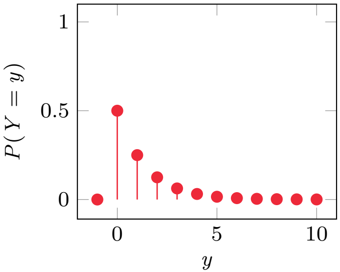
- $g(x) = x - 1 = y$, with inverse $g^{-1}(y) = y + 1$



The range of Y is $\mathcal{D}_Y = \{0, 1, 2, \cdots\}$

Transformations (cont.)

The PMF of Y



$$p_Y(y) = p_X(y+1) = \big(1/2\big)^{y+1}$$
for $y = 0, 1, 2, \dots$

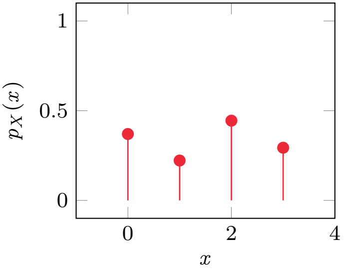
- $p_Y(y) = P(Y = y) = P[g(X) = y] = P[X = g^{-1}(y)] = p_X[g^{-1}(y)]$

Transformations (cont.)

Example

Let the random variable X have the PMF

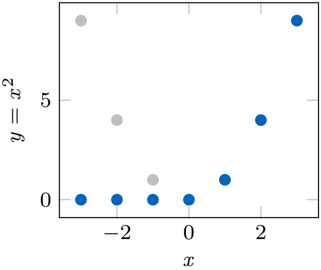
$$p_X(x) = \begin{cases} \dfrac{3!}{x!(3-x!)}(2/3)^x(1/3)^{3-x}, & x = 0,1,2,3 \\ 0, & \text{elsewhere} \end{cases}$$



We are interested in the PMF $p_Y(y)$ of the random variable $Y = X^2$

Transformations (cont.)

$y = g(x) = x^2$ maps $\mathcal{D}_X = \{x : x = 0, 1, 2, 3\}$ onto $\mathcal{D}_Y = \{y : y = 0, 1, 4, 9\}$



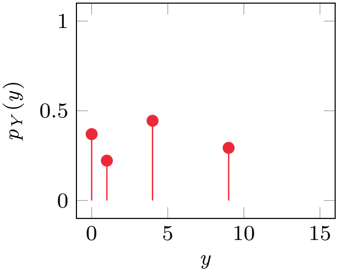
- $y = x^2$ does not always define a one-to-one transformation
- It does here, as the values of x in $\mathcal{D}_X$ are non-negative

Transformations (cont.)

We have a singly valued inverse function $x = g^{-1}(y) = \sqrt{y}$ (not $\pm\sqrt{y}$)

$$p_Y(y) = p_X(\sqrt{y}) = \frac{3!}{(\sqrt{y})!(3-\sqrt{y})!}(2/3)^{\sqrt{y}}(1/3)^{3-\sqrt{y}}$$

$$\text{for } y = 0, 1, 4, 9$$



Transformations (cont.)

g is not one-to-one

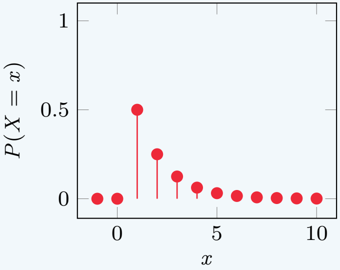
The PMF of discrete Y can be obtained in an uncomplicated manner

- There is no need to develop a rule

Transformations (cont.)

Example

Let $Z = (X-2)^2$, with X the geometric random variable whose PMF is

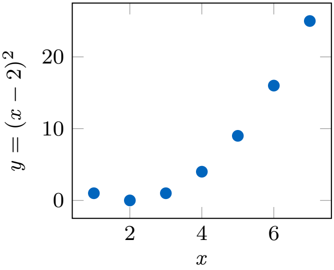


$$P(X = x) = \big(1/2\big)^x$$
$$\text{for } x = 1, 2, 3, \ldots$$

Transformations (cont.)

The range of Z is $\mathcal{D}_Z = \{0, 1, 4, 9, 16, \dots\}$

- $Z = 0$, iff $X = 2$
- $Z = 1$, iff $X = 1$ or $X = 3$

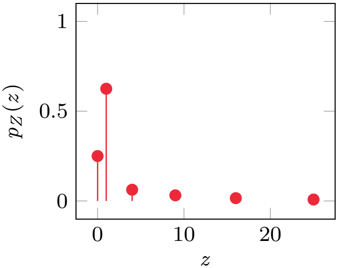


For $z \in \{4, 9, 16, \dots\}$, there is a 1-to-1 map, $x = \sqrt{z} + 2$

Transformations (cont.)

Hence, the PMF of Z is

$$p_Z(z) = \begin{cases} p_X(2) = 1/4, & z = 0 \\ p_X(1) + p_X(3) = 5/8, & z = 1 \\ p_X(\sqrt{z}+2) = 1/4(1/2)^{\sqrt{z}}, & z = 4, 9, 16, \dots \end{cases} \quad (13)$$



Continuous random variables

Random variables

Generalities

Continuous random variables

Continuous random variables

Another class of random variables is the class of RVs of the continuous type

Definition

Continuous random variables

*A random variable is said to be a **continuous random variable** if its cumulative distribution function $F_X(x)$ is a continuous function for all $x \in \mathcal{R}$*

We know that for any random variable X, $P(X = x) = F_X(x) - F_X(x-)$[2]

- Hence, for a continuous RV X, no points can be of discrete mass
- If X is continuous, then it must be $P(X = x) = 0$, for all $x \in \mathcal{R}$

[2] $F_X(x) = \lim_{z \uparrow x} F_X(x)$

Continuous random variables(cont.)

Most common continuous RVs are **absolutely continuous**

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}t, \tag{14}$$

for some function $f_X(t)$

The function $f_X(t)$ is the **probability density function** (**PDF**) of X

If $f_X(t)$ is continuous, then the fundamental theorem of calculus yields

$$\frac{\mathrm{d}}{\mathrm{d}x} F_X(t) = f_X(t) \tag{15}$$

Continuous random variables (cont.)

The **support** of a continuous RV X consists of all points x st $f_X(x) > 0$

- We indicate the support of X by $\mathcal{S}_X$ (as in the discrete case)

If X is a continuous RV, then probabilities are determined by integration

$$P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(t)\mathrm{d}t$$

Moreover, for RVs of the continuous type, we have

$$P(a < X \leq b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X < b)$$

Continuous random variables (cont.)

Note that PDFs satisfy the two properties

- $f_X(x) \geq 0$
- $\int_{-\infty}^{+\infty} f_X(t)\mathrm{d}t = 1$

Second property follows from $F_X(\infty) = 1$

Remark

It can be shown that if a function satisfies the aforementioned properties, such a function is the PDF of a random variable of the continuous type

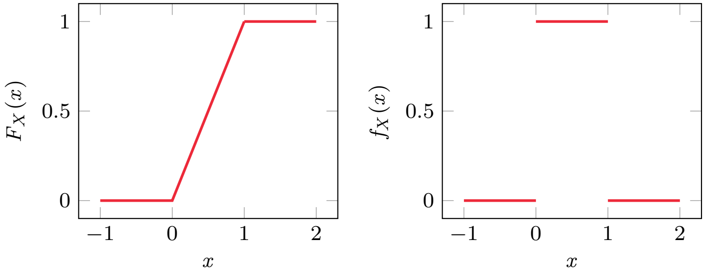
Continuous random variables (cont.)

Example

Pick a real number at random from interval $(0, 1)$, let X be the number

- The chosen number X is an example of a continuous RV
- The space of X is $\mathcal{D} = (0, 1)$

The CDF is $F_X(x) = x$ for $x \in (0, 1)$, its PDF is $f_X(x) = 1$ for $x \in (0, 1)$



Continuous random variables (cont.)

Remark

Continuous/discrete RVs X whose PDF/PMF is constant on support $\mathcal{S}_X$

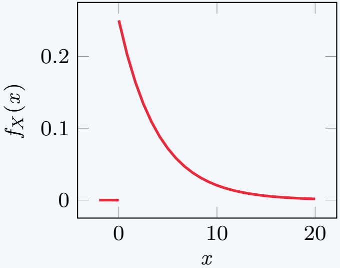
$\rightsquigarrow$ They are said to have a **uniform** distribution

Continuous random variables (cont.)

Example

Let RV X be the time (mins) between incoming chat messages from friends

Suppose that a valid probability model for X is the PDF

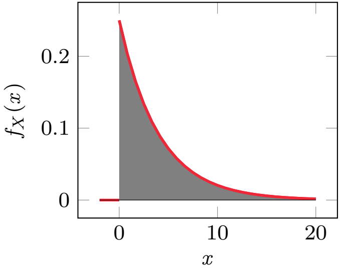


$$f_X(x) = \begin{cases} 0.25e^{(-x/4)}, & 0 < x < \infty \\ 0, & \text{elsewhere} \end{cases}$$

Continuous random variables (cont.)

f_X satisfies the two properties of a PDF

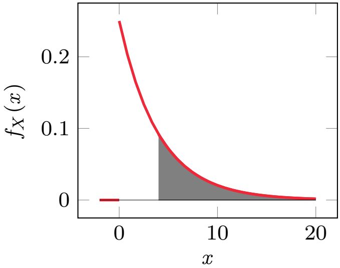
- $f(x) \geq 0$
- $\int_{-\infty}^{+\infty} 0.25 \exp(-x/4) \mathrm{d}x = \Big[-\exp(-x/4) \Big]_0^{\infty} = 1$



Continuous random variables (cont.)

The probability that the time between successive messages is > 4 minutes

$$P(X > 4) = \int_4^{\infty} 0.25 \exp(-x/4) \mathrm{d}x = \exp(-1) \approx 0.368$$



Transformations

Continuous random variables

Transformations

Let X be a random variable of the continuous type with known PDF f_X

We are interested in the distribution of a random variable Y

- Y is some **transformation** of X

⇝ $Y = g(X)$

Often it is possible to determine the PDF of Y by first getting its CDF

Transformations (cont.)

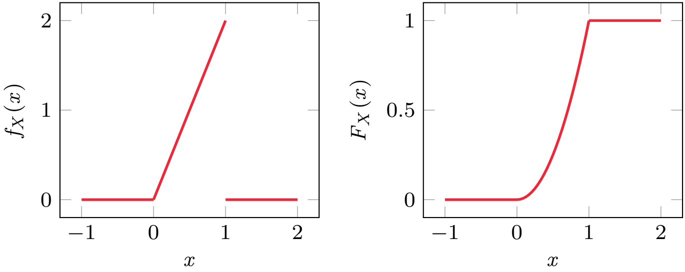
Example

Let X be a random variable whose PDF is given by

$$f_X(x) = \begin{cases} 2x, & 0 \leq x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

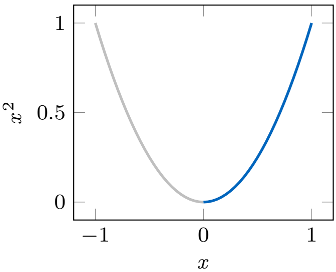
The CDF of X is given by

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \leq x < 1 \\ 1, & 1 \leq x \end{cases} \tag{16}$$



Transformations (cont.)

Suppose we are interested in its square, $Y = X^2$



X and Y have the same support $\mathcal{S}_X = \mathcal{S}_Y = (0, 1)$

- What is the CDF of Y?

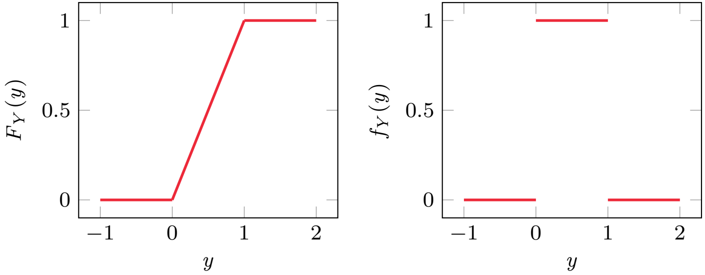
Transformations (cont.)

Using $F_X(x)$ and the fact that $\mathcal{S}_X$ only contains positive numbers

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(X \leq \sqrt{y})$$
$$= F_X(\sqrt{y}) = (\sqrt{y})^2 = y$$

It follows that the PDF of Y is given by

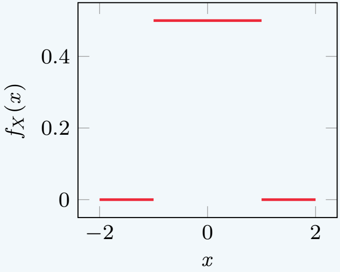
$$f_Y(y) = \begin{cases} 1, & 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$



Transformations (cont.)

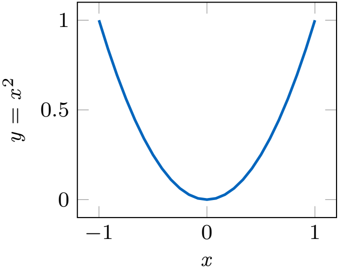
Example

Let $f_X(x) = \dfrac{1}{2}$ for $x \in (-1, 1)$ and zero elsewhere be the PDF of the RV X



Transformations (cont.)

Define the random variable $Y = X^2$

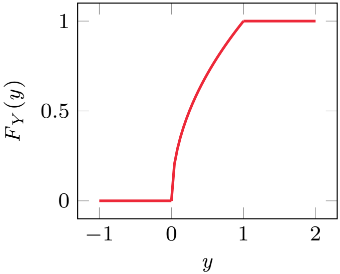


We are interested in the PDF of Y

Transformations (cont.)

If $y \geq 0$, probability $P(Y \leq y)$ equals $P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y})$

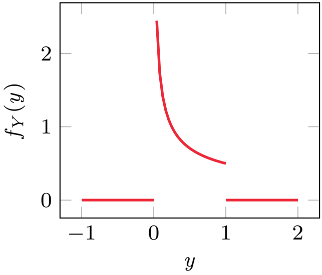
Accordingly, the CDF of Y, $F_Y(y) = P(Y \leq y)$



$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \int_{-\sqrt{y}}^{\sqrt{y}} (\frac{1}{2}) \mathrm{d}x = \sqrt{y}, & y \in [0,1) \\ 1, & 1 \leq y \end{cases}$$

Transformations (cont.)

The PDF of Y



$$f_Y(y) = \begin{cases} \dfrac{1}{2\sqrt{y}}, & 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Transformations (cont.)

We used the **cumulative distribution function technique**

The transformation in the first example is one-to-one

- ⇝ We derive an expression for the PDF of Y
  - In terms of the PDF of X

Transformations (cont.)

Theorem 3.1

Let $f_X(x)$ be the PDF of a continuous random variable X with support $\mathcal{S}_X$

Let $Y = g(X)$, with $g(x)$ a 1-to-1 differentiable function on $\mathcal{S}_X$

Let the inverse of $g(x)$ be denoted by $x = g^{-1}(y)$

Let $dx/dy = d\big[g^{-1}(y)\big]/dy$

Then, the PDF of Y is given by

$$f_Y(y) = f_X\left[g^{-1}(y)\right]\left|\frac{dx}{dy}\right|, \quad \textit{for } y \in \mathcal{S}_Y \tag{17}$$

Set $\mathcal{S}_Y = \{y = g(x) : x \in \mathcal{S}_X\}$ indicates the support of Y

Proof

$g(x)$ is 1-to-1, continuous, strictly monotonically increasing or decreasing

Transformations (cont.)

$g(x)$ is monotonically increasing

The CDF of Y

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P\big[g(X) \leq y\big] = P\big[X \leq g^{-1}(y)\big] \\ &= F_X\left[g^{-1}(y)\right] \end{aligned} \tag{18}$$

The PDF of Y

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_Y(y) = f_X\left[g^{-1}(y)\right] \frac{\mathrm{d}x}{\mathrm{d}y} \tag{19}$$

$\mathrm{d}x/\mathrm{d}y$ is the derivative of function $x = g^{-1}(y)$

As g is increasing ($\mathrm{d}x/\mathrm{d}y > 0$), we can write $\mathrm{d}x/\mathrm{d}y = \big|\mathrm{d}x/\mathrm{d}y\big|$

Transformations (cont.)

$g(x)$ is monotonically decreasing

The CDF of Y

$$F_Y(y) = 1 - F_X\left[g^{-1}(y)\right]$$

The PDF of Y

$$f_Y(y) = f_X\left[g^{-1}(y)\right](-\frac{\mathrm{d}x}{\mathrm{d}y})$$

As g is decreasing ($\mathrm{d}x/\mathrm{d}y < 0$), we can write $-\mathrm{d}x/\mathrm{d}y = \left|\mathrm{d}x/\mathrm{d}y\right|$

Equation (17) is true in both cases

■

Transformations (cont.)

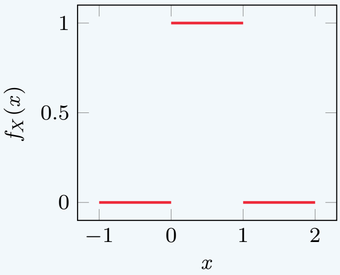
The **Jacobian** of the (inverse) transformation $x = g^{-1}(y)$

$$J = \mathrm{d}x/\mathrm{d}y = \frac{\mathrm{d}[g^{-1}(y)]}{\mathrm{d}y}$$

Transformations (cont.)

Example

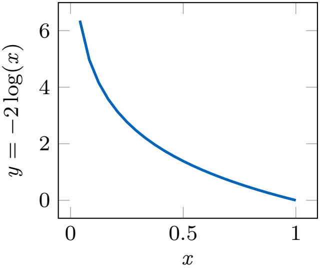
Let the random variable X have the PDF



$$f(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Transformations (cont.)

Consider the random variable $Y = -2\log(X)$

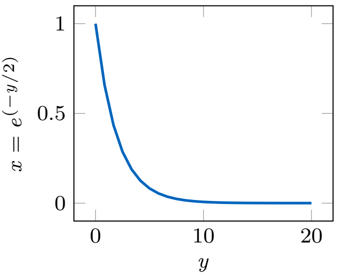


Support sets of X and Y are $\mathcal{S}_X = (0, 1)$ and $\mathcal{S}_Y = (0, \infty)$

The transformation $g(x) = -2\log(x)$ is one-to-one between these sets

Transformations (cont.)

The inverse transformation is $x = g^{-1}(y) = \exp(-y/2)$



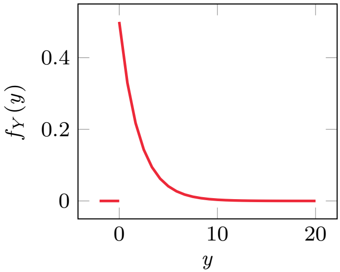
The Jacobian is

$$J = \frac{\mathrm{d}\big[\exp(-y/2)\big]}{\mathrm{d}y} = -1/2 \exp(-y/2)$$

Transformations (cont.)

Accordingly, the PDF of $Y = -2\log(X)$ is

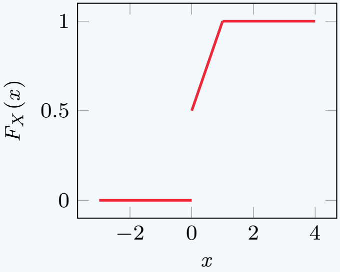
$$f_Y(y) = \begin{cases} f_X\Big[\exp(-y/2)\Big]|J| = 1/2\exp(-y/2), & 0 < y < \infty \\ 0, & \text{elsewhere} \end{cases}$$



Transformations (cont.)

Example

Consider the following distribution function



$$F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x+1}{2}, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

- The distribution is neither of the continuous nor of the discrete type
- $F(x)$ is not always continuous, nor is it a step function

Then,

$$P(-3 < X \leq 1/2) = F(1/2) - F(-3) = 3/4 - 0 = 3/4$$

$$P(X = 0) = F(0) - F(0-) = 1/2 - 0 = 1/2$$

Random variable (cont.)

Distributions that are mixtures of continuous/discrete type are frequent

Example

In survival analysis, we know that life duration X exceeds some number a

- The exact value of X is however unknown (**censoring**)

A classic: A subject under study at some point a disappears

- We know that the subject has lived a certain time a
- The exact life duration of life is unknown