Chapter 2 Review Random Variables

Formally, a random variable is a measurable function from a sample space Ω with a probability measure into $\mathbb{R}$.

i.e. Let Ω be a sample space with probability measure P, then a random variable X, is a function

$$X : \Omega \to \mathbb{R}$$

Example: Toss a biased coin twice $P(\text{Heads}) = \frac{1}{4}$

$$\Omega = \{HH, HT, TH, TT\}$$

prob: $\frac{1}{16}$ $\frac{3}{16}$ $\frac{3}{16}$ $\frac{9}{16}$

$$X : \Omega \to \mathbb{R}$$

X = # of heads

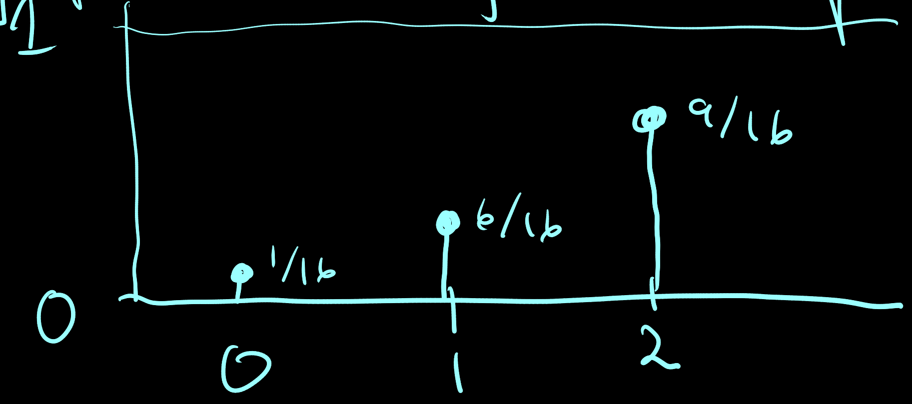
$$\Pr(X = 0) = \Pr(\omega \in \Omega : X(\omega) = 0)$$
$$= P(X^{-1}(0))$$
$$= P(\{HH\}) = 1/16$$

$$\Pr(X=1) = P(X^{-1}(1)) = P(\{HT, TH\}) = 6/16$$

$$\Pr(X=2) = 9/16$$

Often, Ω is not explicit and we just consider P on subsets of $\mathbb{R}$.

$\boxed{\text{PMF}}$: probability mass function : discrete RVs



$$p(x) = \Pr(X=x)$$

$$\sum_x p(x) = 1$$

$\boxed{\text{CDF}}$: Cumulative distribution function

$$F(x) = P(X \leq x) \qquad x \in \mathbb{R}$$

<u>What does F look like ?</u>

$$F : \mathbb{R} \to [0, 1]$$

F is mononotone non-decreasing

$$x_1 \leq x_2 \implies F(x_1) \leq F(x_2$$

Two useful results:

$\boxed{P \text{ is "continuous"}}$

Thm: Let $A_1, A_2, \ldots$ be a monotone increasing sequence of sets (i.e. $A_1 \subset A_2 \subset A_3 \subset \ldots$)

note: includes =

and let $A = \bigcup_{i=1}^{\infty} A_i$

Then $P(A) = \lim_{i\to\infty} P(A_i)$

Proof: Let $D_1 = A_1$, $D_i = A_i - A_{i-1}$ for $i > 1$

Then $A = \bigcup_{i=1}^{\infty} D_i$ and D_i's are disjoint so $\quad P(A) = \sum_{i=1}^{\infty} P(D_i)$ $\quad$ (σ-additivity)

But $P(A_i) = \sum_{j=1}^{i} P(D_i)$

which is a partial sum of

$\sum_{i=1}^{\infty} P(D_i)$

So $\lim_{j\to\infty} P(A_j) = \lim_{j\to\infty} \sum_{i=1}^{j} P(D_i)$

$$= \sum_{i=1}^{\infty} P(D_i)$$

$$= P(A)$$

Q.E.D.

Continuity Theorem Part 2

Let $A_1 \supset A_2 \supset \dots$ be a monotone decreasing sequence of sets.

Let $A = \bigcap_{i=1}^{\infty} A_i$.

Then $\lim_{i \to \infty} P(A_i) = P(A)$

Proof: Consider that $A_1^c \subset A_2^c \subset \dots$ is monotone increasing and apply the previous theorem. [Complete the details]

These theorems are useful to understand the continuity – or lack thereof – of CDFs.

EXERCISES

1) Let $a < 1$, $b > 4$.

a) Find

$$\bigcup_{i=1}^{\infty} \left(a+\frac{1}{n}, b-\frac{1}{n}\right]$$

$$\bigcap_{i=1}^{\infty} \left(a-\frac{1}{n}, b+\frac{1}{n}\right]$$

$$\bigcup_{i=1}^{\infty} \left[a+\frac{1}{n}, b-\frac{1}{n}\right)$$

$$\bigcap_{i=1}^{\infty} \left[a-\frac{1}{n}, b+\frac{1}{n}\right)$$

Show that F is <u>right continuous</u>

i.e. $\lim_{x \to c^+} F(x) = F(c) \quad \forall c \in \mathbb{R}.$

i.e. if $x_1 > x_2 > x_3 \cdots\cdots > c$
and $\lim_{i \to \infty} x_i = c$

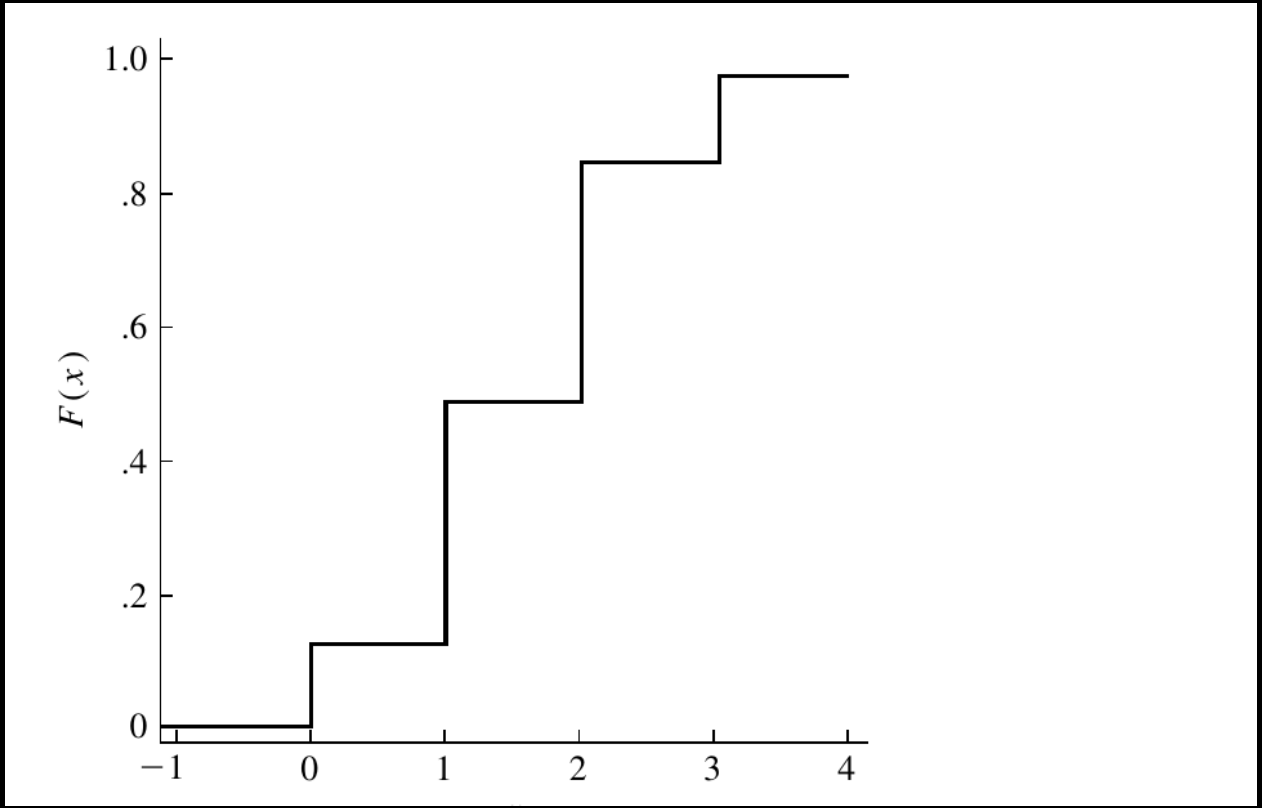
then $\lim_{i \to \infty} F(x_i) = F(c)$

Hint: Consider the fact that

$$F(x_i) = P_X\big((-\infty, x_i]\big)$$

and apply the continuity theorem.

When you draw a CDF show right continuity.
A CDF for a discrete distribution does not
look like:



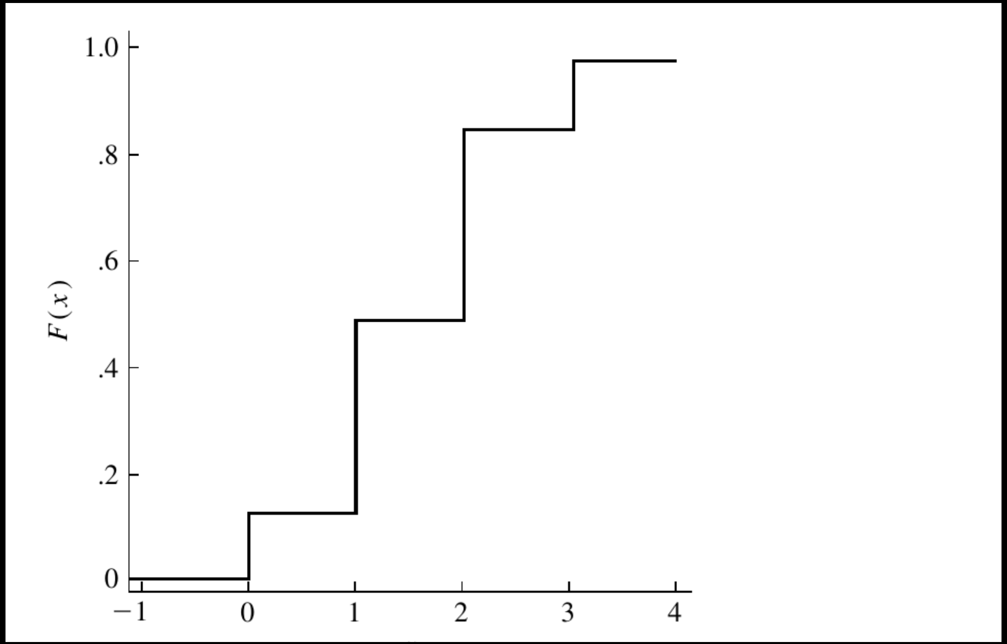
This is a CDF
in the text for X
with PMF

$p(0) = 1/8$

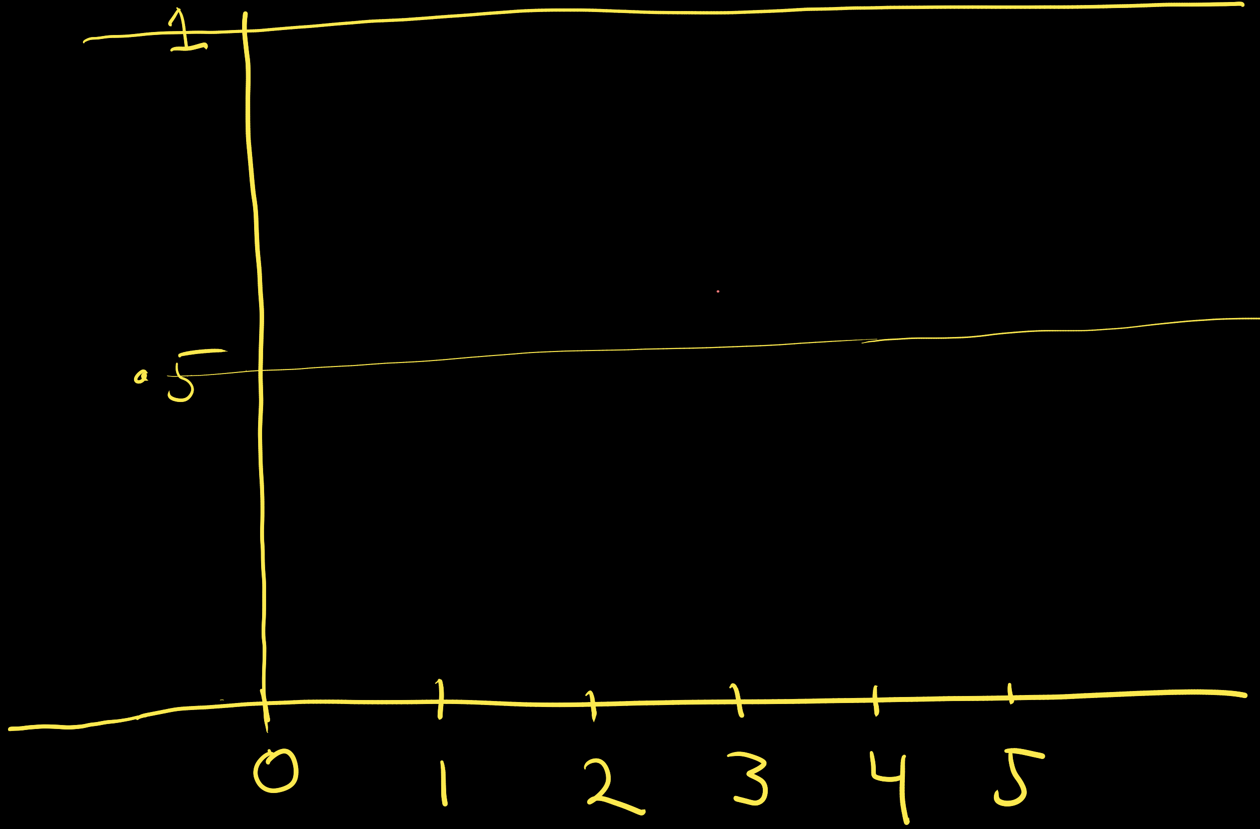
$p(1) = 3/8$

$p(2) = 3/8$

$p(3) = 1/8$



1.0
.8
.6
.4
.2
0
F(x)
−1
0
1
2
3
4



1
.5
0
1
2
3
4
5

More on CDFs

1) $\lim_{x \to -\infty} F(x) = 0$ Proof: Use continuity thm

2) $\lim_{x \to +\infty} F(x) = 1$ " "

3) $P(X \in (a, b]) = F(b) - F(a)$

4) If $P(X = x) = p$ then $F(x) - \underbrace{\lim_{n \to \infty} F(x - \frac{1}{n})}_{\lim_{x' \to x^-} F(x')} = p$

"limit from the left"

Definition (not in text) "support"

For a discrete R.V., X,

$$\operatorname{supp}(X) = \{x : p(x) > 0\}$$

In general, if f is a function $\to \mathbb{R}$, then the $\operatorname{supp}(f) = \bigcap$ closed sets containing $\{x : f(x) \neq 0\}$

More later. For now, 2 discrete RVs are <u>statistically independent</u> if:

$$P(X = x \text{ and } Y = y) = P(X = x) \times P(Y = y)$$

$$\forall\, x \in \text{supp}(X),\ y \in \text{supp}(Y)$$

If X, Y, Z are discrete RVs then X, Y, Z are mutually independent.

$$P(X=x, Y=y, Z=z) = P(X=x) \times P(Y=y) \times P(Z=z)$$

$$\forall x \in supp(X),\ y \in supp(Y),\ z \in supp(Z)$$

<u>note:</u> Mutually independent $\Longrightarrow$ pairwise independent

<u>BUT</u> Does it go the other way?

CAUTION: Let A, B, C be events.

Is it possible to have A & B independent

A & C "

B & C "

But A, B, C not mutually independent

I.E. $P(A \cap B \cap C) \neq P(A)P(B)P(C)$

Proof or counterexample ??

Bernoulli R.V.s
0-1 R.V.s

Toss a biased coin. $P(H)=p$

$X = \#$ of heads

$$p(1) = p \qquad p(0) = 1-p = q$$

OR: $$p(x) = \begin{cases} p^{x}(1-p)^{1-x} & \text{if } x \in \{0,1\} \\ 0 & \text{otherwise} \end{cases}$$

Useful Idea: Indicator variable.

Event A. Let $$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \omega \notin A. \end{cases}$$

I_A is a Bernoulli R.V with $P(I_A = 1) = P(A)$

Binomial Distribution :

If $X_1, X_2, \dots, X_n$ are independent Bern.(p)

Then $Y = \sum_{i=1}^{n} X_i$ is Binomial (n, P)

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$k = 0, 1, \dots n$

Let $X_1, X_2, \ldots, X_k, X_{k+1}, \ldots$ be Bern(p)

Geometric : X be # of 0's & 1's until first 1. e.g.

negative binomial : Y " " " 0's & 1's until r_{th} 1

Bern X_is : $\underbrace{0 \quad 0 \quad 1}_{3} \quad 0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad \ldots$

Geometric : X 3

neg. Bin X_3 8

$r = 3$

What is neg. bin. X for $r = 2$?

Bern. X_i's: 1 0 1 1 0 0 1 0 1 ...

Geom X 1

0

Neg. Bin

$r = 1$ 1

$r = 3$ 4

PMFs

Geometric: $p(k) = (1-p)^{k-1} p \qquad k = 1, 2, \dots,$

Neg Bin r: $p(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$

Property: Let $X_1, \ldots, X_r$ be r ind. Bern(p)

Let $Y = X_1 + X_2 + \cdots + X_r$

Then $Y \sim \text{NegBin}(r, p)$

Binomial : Bernoulli : :

NegBin : ________?________

If $X_1, \dots, X_n$ are independent Bern(p)

then $X = X_1 + X_2 + \dots X_n$

$\sim$ ______________ ?

If $X_1, \dots, X_n$ are independent Geometric(p)

then $X_1 + X_2 + \dots + X_n \sim$ ______________ ?

Hypergeometric - Review

Poisson (λ) Example: # of fatal car accidents in Toronto in a day

$$\boxed{X = 0, 1, 2, \ldots}$$

λ = average number/day

Consider Maclaurin series for e^{λ}

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{6} + \frac{\lambda^4}{24} + \ldots + \frac{\lambda^k}{k!} + \ldots$$

Def $X \sim \text{Poisson}(\lambda)$, $\lambda > 0$

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Exer: Show $p(x)$ is a PMF.

1) $p(x) \geq 0$ since $e^{-\lambda} > 0$, $\lambda^x > 0$ for $\lambda > 0$

and $\frac{1}{x!}$ is ________________?

What about $x = 0$?

2) $$\sum_{x=0}^{\infty} p(x) = \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

The Poisson as a limit of binomials

Take a very rare event that happens randomly by chance for each person BUT you have a large population.

E.G. Winning 6-49.

For each ticket the probability is

$$\frac{1}{\binom{49}{6}} =$$

```
> 1/choose(49,6)
[1] 7.151124e-08
```

$= 0.0000000715\ldots$

Suppose n people buy tickets

AND

suppose they all choose #'s randomly

false assumption

Let X = # of winners.

What is $P(X=0)$? $P(X>1)$?

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

if p and k fixed "small"

and n very "large"

Use asymptotics : Let $n \to \infty$ but p small

Let $\lambda = np$ so $p = \frac{\lambda}{n}$

$$p(k) = \frac{n!}{(n-k)!\,k!} \left(\frac{\lambda}{n}\right)^{k} \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

Use asymptotics : Let $n \to \infty$ but p small

Let $\lambda = np$ so $p = \frac{\lambda}{n}$

$$p(k) = \frac{n!}{(n-k)!\,k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

look like Poisson – missing $e^{-\lambda}$

$$= \frac{\lambda^k}{k!}\underbrace{\frac{n!}{(n-k)!}\frac{1}{n^k}}_{\frac{n\times(n-1)\times\cdots\times(n-k+1)}{n\times n\times\cdots\times n}}\underbrace{\left(1-\frac{\lambda}{n}\right)^{n}}\left(1-\frac{\lambda}{n}\right)^{-k}$$

$\underbrace{n\times n\times\cdots\times n} \to \boxed{}$?

$\to \boxed{}$? $\to \boxed{}$?

So X is Poisson (λ) $\lambda = np$

Of 28,000,000 buy tickets $\lambda \approx \frac{28\,000\,000}{14\,000\,000} = 2$

So

$$P(X=0) = \frac{e^{-\lambda}\lambda^0}{0!} = \frac{e^{-2}2^0}{0!} = 0.1353$$

$$P(X=1) = \frac{e^{-\lambda}2^1}{1!} = \frac{e^{-2}2}{1!} = 0.2706$$

$$P(X>1) = 1 - 0.1353 - 0.2706 = 0.5941$$

Why are assumptions incorect and what's the consequence?

Continuous random variables

Probability by integrating over a density: PDF

$$P(a < X < b) = \int_a^b f(x)\,dx$$

Requirements for f:

1) $f(x) \geq 0$ almost everywhere

2) $$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

Important examples

Uniform$(0,1)$ $\quad U(0,1)$

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

Uniform(a, b) with $a < b$

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

Exponential (λ) $\quad \lambda > 0$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

- Lifetime with no aging.
- Distribution of lifetime if only cause of death is accidents.
- "memoryless"

Gamma density

$\alpha > 0, \lambda > 0$

$$f(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \quad t > 0$$

where $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} \, du, \quad \alpha > 0$

$$\Gamma(1) = 1$$

$$\Gamma(n+1) = n\Gamma(n)$$

uses integration by parts

By induction $\Gamma(n+1) = n!$ for $n = 0, 1, 2, \ldots$

Normal $\quad Z \sim N(0,1)$ Standard Normal

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad z \in \mathbb{R}$$

Beta

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

$$0 < x < 1$$

Functions of a random variable

Linear transformations.

Let X have PDF f and CDF F.
continuous

Let $Y = aX + b$, $b \in \mathbb{R}$, $a > 0$

Then $F_Y(y) = P(Y \le y)$

$= P(aX + b \le y)$

$= P(aX \le y - b)$

$= P\left(X \le \frac{y-b}{a}\right)$

$$= F_X\left(\frac{y-b}{a}\right)$$

So PDF of Y is

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

$$= \frac{d}{dy} F_X\left(\frac{y-b}{a}\right)$$

$$= f_X\left(\frac{y-b}{a}\right) \times \frac{1}{a}$$

by chain rule

Normal density:

$$Y = \underset{\text{scale}}{\sigma} Z + \underset{\text{location}}{\mu} \qquad X \sim N(0,1)$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$$

$$y \in \mathbb{R}$$

Change of variable formula

X continuous with density f

$Y = g(X)$, g differentiable, strictly mononotone on $\operatorname{supp}(X)$

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

and $f_Y(y) = f_X(g^{-1}(y)) \times \frac{d}{dy} g^{-1}(y)$