Subsection 5

Functions of a continuous random variable

Why functions?

- We may model one quantity as a random variable X. We may have to work with another closely related quantity
- Example 1
  - Length of a square: X
  - Area of the square: $Y = X^2$
- Example 2
  - Volume of a liquid: X
  - Density: ρ
  - Weight ~~Volume occupied~~: $Y = \rho X$
- Given the distribution of X, it is useful to have a method for finding the distribution of a function of X

Example

Suppose $X \sim$ Uniform$[0, 1]$

- $Y = 2X \in [0, 2]$ is clearly a random variable
- What is the distribution of Y?

Example

Suppose $X \sim$ Uniform$[0,1]$

- $Y = 2X \in [0,2]$ is clearly a random variable
- What is the distribution of Y?

For $y \in [0,2]$, (Find CDF first)

$$F_Y(y) = P(Y \leq y) = P(2X \leq y)$$

$$= P(X \leq y/2) = \int_0^{\frac{y}{2}} \underbrace{f_X(x)}_{=1} dx = \frac{y}{2}.$$

Example

Suppose $X \sim$ Uniform$[0, 1]$

- $Y = 2X \in [0, 2]$ is clearly a random variable
- What is the distribution of Y?

For $y \in [0, 2]$,

$$F_Y(y) = P(Y \leq y) = P(2X \leq y)$$

$$= P(X \leq y/2) = \int_0^{\frac{y}{2}} f_X(x)dx = \frac{y}{2}.$$

PDF of Y, $f_Y(y) = \frac{dF_Y(Y)}{dy} = \frac{1}{2}$.

$$Y \sim \text{Uniform}[0, 2]$$

Example

$Y = X+5$? (or) $Y = 2X+5$?

$Y = aX+b \sim$ Uniform$[b, b+a]$

Suppose $X \sim$ Uniform$[0, 1]$

- $Y = 2X \in [0, 2]$ is clearly a random variable
- What is the distribution of Y?

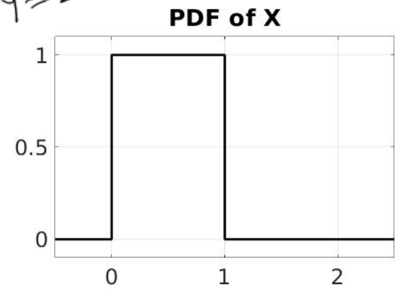
For $y \in [0, 2]$,

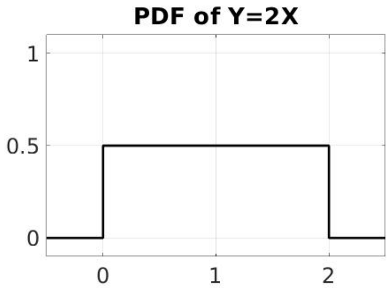
$$F_Y(y) = P(Y \le y) = P(2X \le y)$$

$$= P(X \le y/2) = \int_0^{\frac{y}{2}} f_X(x)dx = \frac{y}{2}.$$

PDF of Y, $f_Y(y) = \frac{dF_Y(Y)}{dy} = \frac{1}{2}$.

$$Y \sim \text{Uniform}[0, 2]$$





General case: CDF of $g(X)$

- Suppose X is a continuous random variable with CDF F_X and PDF f_X

General case: CDF of $g(X)$

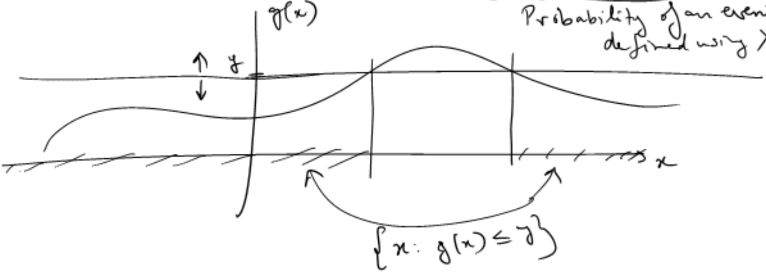
- Suppose X is a continuous random variable with CDF F_X and PDF f_X
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a (reasonable) function

General case: CDF of $g(X)$

- Suppose X is a continuous random variable with CDF F_X and PDF f_X
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a (reasonable) function
- Then, $Y = g(X)$ is a random variable with CDF F_Y determined as follows:

General case: CDF of $g(X)$

- Suppose X is a continuous random variable with CDF F_X and PDF f_X
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a (reasonable) function
- Then, $Y = g(X)$ is a random variable with CDF F_Y determined as follows:
  - $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \underbrace{P(X \in \{x : g(x) \leq y\})}_{\text{Probability of an event defined using } X}$



General case: CDF of $g(X)$

- Suppose X is a continuous random variable with CDF F_X and PDF f_X
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a (reasonable) function
- Then, $Y = g(X)$ is a random variable with CDF F_Y determined as follows:
  - $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \in \{x : g(x) \leq y\})$
- How to evaluate the above probability?

General case: CDF of $g(X)$

- Suppose X is a continuous random variable with CDF F_X and PDF f_X
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a (reasonable) function
- Then, $Y = g(X)$ is a random variable with CDF F_Y determined as follows:
  - $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \in \{x : g(x) \leq y\})$
- How to evaluate the above probability?
  - Convert the subset $A_y = \{x : g(x) \leq y\}$ into intervals in real line

General case: CDF of $g(X)$

- Suppose X is a continuous random variable with CDF F_X and PDF f_X
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a (reasonable) function
- Then, $Y = g(X)$ is a random variable with CDF F_Y determined as follows:
  - $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \in \{x : g(x) \leq y\})$
- How to evaluate the above probability?
  - Convert the subset $A_y = \{x : g(x) \leq y\}$ into intervals in real line
  - Find the probability that X falls in those intervals

General case: CDF of $g(X)$

- Suppose X is a continuous random variable with CDF F_X and PDF f_X
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a (reasonable) function
- Then, $Y = g(X)$ is a random variable with CDF F_Y determined as follows:
  - $F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \in \{x : g(x) \le y\})$
- How to evaluate the above probability?
  - Convert the subset $A_y = \{x : g(x) \le y\}$ into intervals in real line
  - Find the probability that X falls in those intervals
  - $F_Y(y) = P(X \in A_y) = \int_{A_y} f_X(x)dx$

General case: CDF of $g(X)$

- Suppose X is a continuous random variable with CDF F_X and PDF f_X
- Suppose $g : \mathbb{R} \to \mathbb{R}$ is a (reasonable) function
- Then, $Y = g(X)$ is a random variable with CDF F_Y determined as follows:
  - $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \in \{x : g(x) \leq y\})$
- How to evaluate the above probability?
  - Convert the subset $A_y = \{x : g(x) \leq y\}$ into intervals in real line
  - Find the probability that X falls in those intervals
  - $F_Y(y) = P(X \in A_y) = \int_{A_y} f_X(x)dx$
- If F_Y has no jumps, you may be able to differentiate and find a PDF

Monotonic, differentiable functions

Theorem

Suppose X is a continuous random variable with PDF f_X. Let $g(x)$ be monotonic for $x \in supp(X)$ with derivative $g'(x) = \frac{dg(x)}{dx}$. Then, the PDF of $Y = g(X)$ is

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

$g'(x)$

$g^{-1}(y)$

$f_X(x)$

$g^{-1}(y)$

Monotonic, differentiable functions

Theorem

Suppose X is a continuous random variable with PDF f_X. Let $g(x)$ be monotonic for $x \in supp(x)$ with derivative $g'(x) = \frac{dg(x)}{dx}$. Then, the PDF of $Y = g(X)$ is

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

- Translation: $Y = X + a$ $\quad g(x) = x + a, \quad g'(x) = 1, \quad y = x + a$

$x = y - a$

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

$$f_Y(y) = f_X(y - a)$$

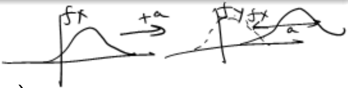
Monotonic, differentiable functions

Theorem

Suppose X is a continuous random variable with PDF f_X. Let $g(x)$ be monotonic for $x \in supp(X)$ with derivative $g'(x) = \frac{dg(x)}{dx}$. Then, the PDF of $Y = g(X)$ is

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

- Translation: $Y = X + a$



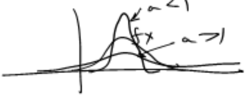
$$f_Y(y) = f_X(y - a)$$

- Scaling: $Y = aX$

$g(x) = ax$, $g'(x) = a$, $y = ax \Rightarrow x = y/a$

$g^{-1}(y) = y/a$

$$f_Y(y) = \frac{1}{|a|} f_X(y/a)$$



Monotonic, differentiable functions

Theorem

Suppose X is a continuous random variable with PDF f_X. Let $g(x)$ be monotonic for $x \in supp(X)$ with derivative $g'(x) = \frac{dg(x)}{dx}$. Then, the PDF of $Y = g(X)$ is

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

- Translation: $Y = X + a$

$$f_Y(y) = f_X(y - a)$$

- Scaling: $Y = aX$

$$f_Y(y) = \frac{1}{|a|} f_X(y/a)$$

- Affine: $Y = aX + b$

$$f_Y(y) = \frac{1}{|a|} f_X((y - b)/a)$$

Affine transformation of normal distributions

- $X \sim$ Normal$(0,1)$ Standard normal

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$$

replace x with $\frac{y-\mu}{\sigma}$

  - $Y = \sigma X + \mu$ "Affine" $(\sigma > 0)$

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-(y-\mu)^2/2\sigma^2)$$: PDF of Normal(μ, σ^2)

$Y \sim$ Normal(μ, σ^2)

- $X \sim$ Normal(μ, σ^2) $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ (Exercise)

replace x with $\frac{y + \mu/\sigma}{1/\sigma} = \sigma y + \mu$

  - $Y = (X-\mu)/\sigma \sim$ Normal$(0,1)$

affine

$Y = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$

Affine transformation of normal distributions

- $X \sim$ Normal$(0, 1)$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$$

  - $Y = \sigma X + \mu$

$$f_Y(\overset{y}{\cancel{x}}) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-(\overset{y}{\cancel{x}} - \mu)^2/2\sigma^2)$$

  $Y \sim$ Normal(μ, σ^2)

- $X \sim$ Normal(μ, σ^2)
  - $Y = (X - \mu)/\sigma \ \sim$ Normal$(0, 1)$

Result

Affine transformation of a normal random variable is normal.

Problem

Let $X \sim \text{Exp}(\lambda)$. Find the PDF of X^2.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

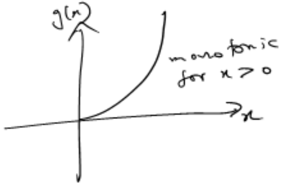
$$\text{supp}(X) = \{x : x > 0\}$$

$g(x) = x^2$

$g'(x) = 2x$

$y = x^2 \Rightarrow x = \sqrt{y}$

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



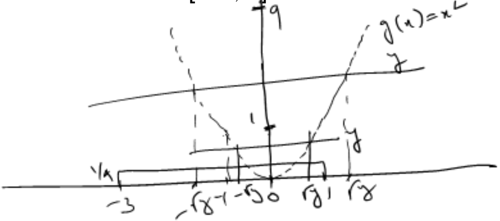
$Y = X^2$

$$f_Y(y) = \frac{1}{2\sqrt{y}} \lambda e^{-\lambda\sqrt{y}}, \quad y > 0$$

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

Problem

Let $X \sim \text{Uniform}[-3,1]$. Find the PDF of X^2.



$Y = X^2 \in [0, 9]$

$\text{supp}(X) = [-3, 1]$ $\quad g(x) = x^2$ is not monotonic in $\text{supp}(X)$

$y \in [0, 9)$ $\quad F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$

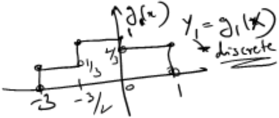
$\underline{y \in [0,1]}$: $(X^2 \leq y) \Longleftrightarrow -\sqrt{y} < X < \sqrt{y} \Rightarrow F_Y(y) = \frac{2\sqrt{y}}{4}$

$y \in [1, 9]$: $(X^2 \leq y) \Longleftrightarrow -\sqrt{y} < X < 1 \Rightarrow F_Y(y) = \frac{1+\sqrt{y}}{4}$

$$f_Y(y) = \begin{cases} \frac{1}{2} \cdot \frac{1}{2\sqrt{y}}, & 0 < y < 1 \\ \frac{1}{4} \cdot \frac{1}{2\sqrt{y}}, & 1 < y < 9 \end{cases}$$

Problem

Let $X \sim \text{Uniform}[-3, 1]$. Find the PDF of $\max(X, 0)$.

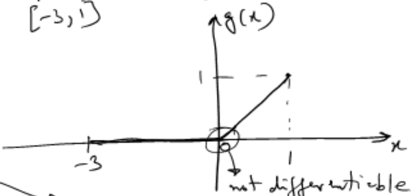


$Y = g(X) \in [0, 1]$

$g(x) = \max(x, 0) = \begin{cases} 0, & \text{if } -3 < x < 0 \\ x, & \text{if } 0 < x < 1 \end{cases}$

$x \in \text{Supp}(X)$
$[-3, 1]$

$y < 0$: $F_Y(y) = P(Y \le \text{negative number}) = 0$



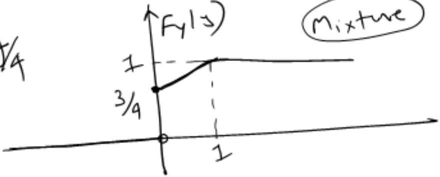
$y = 0$:
$F_Y(y) = P(Y \le 0) = P(Y = 0)$
$= P(g(X) = 0)$
$= P(-3 < X < 0) = 3/4$

Y: not continuous

Mixture

$0 < y < 1$: $F_Y(y) = \frac{3}{4} + y/4$
(exercise)

$y > 1$: $F_Y(y) = 1$



Subsection 6

Continuous random variables: Expected value

Expected value: Function of a continuous random variable

Theorem

Let X be a continuous random variable with density $f_X(x)$. Let $g : \mathbb{R} \to \mathbb{R}$ be a function. The expected value of $g(X)$, denoted $E[g(X)]$, is given by

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

whenever the above integral exists.

Expected value: Function of a continuous random variable

Theorem

Let X be a continuous random variable with density $f_X(x)$. Let $g : \mathbb{R} \to \mathbb{R}$ be a function. The expected value of $g(X)$, denoted $E[g(X)]$, is given by

$$E[g(X)] = \underbrace{\int_{-\infty}^{\infty}}_{\text{supp}(X)} g(x)\underbrace{f_X(x)dx}_{\text{like PMF}},$$

whenever the above integral exists.

- If X is discrete with range T_X and PMF p_X,

$$E[g(X)] = \sum_{x \in T_X} g(x) p_X(x)$$

  - **Summation in discrete case is replaced by integration in coninuous case**

- The integral may diverge to $\pm\infty$ or may not exist in some cases

Mean and Variance

X: continuous random variable

- Mean, denoted $E[X]$ or μ_X or simply μ

$g(x) = x$

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

  - Mean is the average or expected value of X

- Variance, denoted $\text{Var}(X)$ or σ_X^2 or simply σ^2

$$\text{Var}(X) = E[\underbrace{(X - \mu_X)^2}_{g(x) = (x - \mu_x)^2}] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)dx$$

  - Variance is a measure of spread of X about its mean
  - $\text{Var}(X) = E[X^2] - E[X]^2$

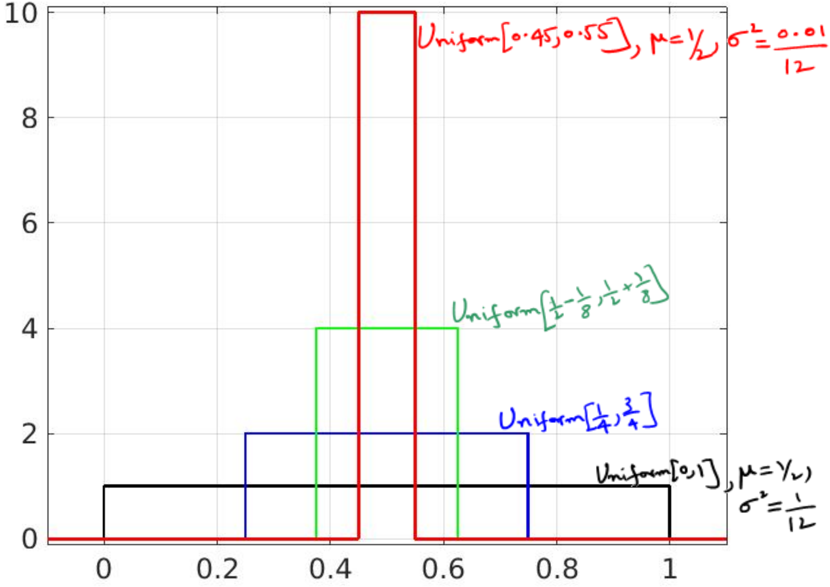
- Evaluating expected value needs good knowledge of integration
  - Formulae are available in numerous webpages and books

Examples of mean and variance

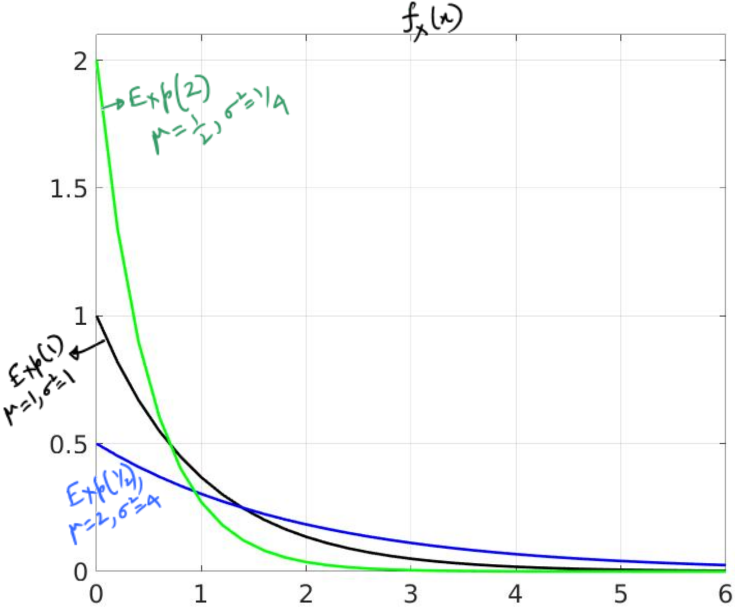
- $X \sim$ Uniform$[a, b]$, $f_X(x) = \dfrac{1}{b-a}$, $a < x < b$
  - $E[X] = \dfrac{a+b}{2}$, $\text{Var}(X) = \dfrac{(b-a)^2}{12}$
- $X \sim$ Exp(λ), $f_X(x) = \lambda \exp(-\lambda x)$, $x > 0$
  - $E[X] = 1/\lambda$, $\text{Var}(X) = 1/\lambda^2$
- $X \sim$ Normal(μ, σ^2), $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-(x-\mu)^2/2\sigma^2)$
  - $E[X] = \mu$, $\text{Var}(X) = \sigma^2$

$$\left.\begin{aligned} &\int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu \\ &\int_{-\infty}^{\infty} (x-\mu)^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma^2 \end{aligned}\right\} \text{more involved integration}$$

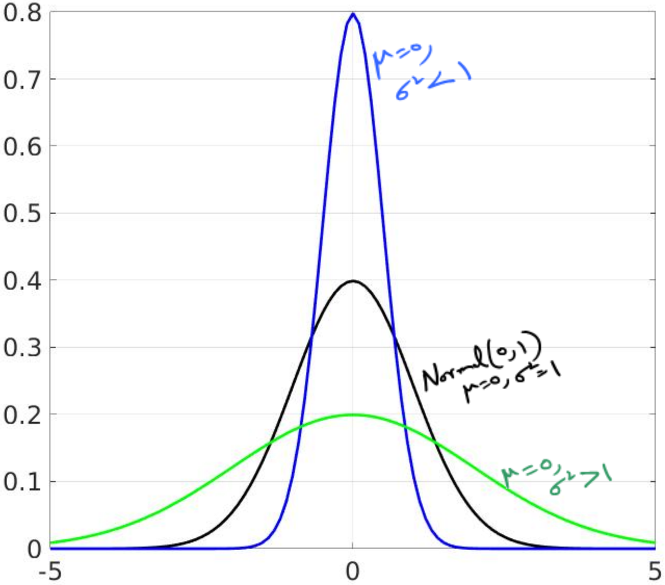
Uniform distribution with different variances



Exponential distribution with different λ



Normal distribution with different σ



Markov and Chebyshev inequalities

- Markov inequality
  - X: continuous random variable with mean μ
  - supp(X): non-negative, i.e. $P(X < 0) = 0$

$$P(X > c) \leq \frac{\mu}{c}$$

- Chebyshev inequality
  - X: continuous random variable with mean μ and variance σ^2

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

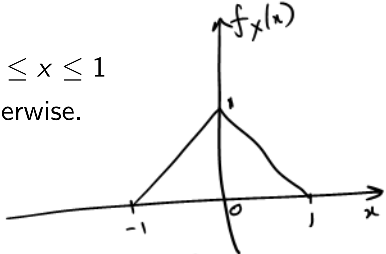
Probability space and its axioms

- Discrete case
  - Sample space: finite or countable set
  - Events: power set of sample space
  - Probability function: PMF
- Continuous case
  - Sample space: interval of real line
  - Events: intervals in the sample space along with their complements and countable unions
    - ★ This avoids some 'bizarre' subsets that defy our sense of measure
  - Probability function: function from intervals inside sample space to $[0, 1]$ satisfying the axioms
    - ★ Possible only if $P(X = x) = 0$
- Unified description of probability spaces: Measure-theoretic
  - NPTEL course: https://nptel.ac.in/courses/108/106/108106083/

Problem

A continuous random variable X has PDF

$$f_X(x) = \begin{cases} 1 - |x|, & -1 \le x \le 1 \\ 0, & \text{otherwise.} \end{cases}$$



Find the CDF of X, $E[X]$, $\mathrm{Var}(X)$.

$$F_X(x) = \begin{cases} 0, & x < -1 \\ ?, & -1 \le x < 0 \\ ?, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

$E[X] = 0$

$-1 \le x \le 0$

$$F_X(x) = \underbrace{\int_{-1}^{x} (1 - |u|)\, du}_{x=0} = \int_{-1}^{x} (1+u)\, du = u\Big|_{-1}^{x} + \frac{u^2}{2}\Big|_{-1}^{x} = (x - (-1)) + \left(\frac{x^2}{2} - \frac{(-1)^2}{2}\right)$$

$$= x + 1 + \frac{x^2}{2} - \frac{1}{2} = \frac{1}{2} + x + \frac{x^2}{2}$$

$0 \le x \le 1$

$$F_X(x) = \underbrace{\int_{-1}^{0} (1 - |u|)\, du}_{1/2} + \int_{0}^{x} (1 - |u|)\, du = \frac{1}{2} + u\Big|_{0}^{x} - \frac{u^2}{2}\Big|_{0}^{x} = \frac{1}{2} + x - \frac{x^2}{2}$$

$$E[X] = \int_{-1}^{1} x f_X(x)\,dx = \int_{-1}^{0} x(1+x)\,dx + \int_{0}^{1} x(1-x)\,dx$$

$$= \frac{x^2}{2}\Big|_{-1}^{0} + \frac{x^3}{3}\Big|_{-1}^{0} + \frac{x^2}{2}\Big|_{0}^{1} - \frac{x^3}{3}\Big|_{0}^{1} = -\frac{1}{2} + \frac{1}{3} + \frac{1}{2} - \frac{1}{3} = 0$$

$$\mathrm{Var}(X) = E[X^2] = \int_{-1}^{0} x^2(1+x)\,dx + \int_{0}^{1} x^2(1-x)\,dx$$

$$= \frac{x^3}{3}\Big|_{-1}^{0} + \frac{x^4}{4}\Big|_{-1}^{0} + \frac{x^3}{3}\Big|_{0}^{1} - \frac{x^4}{4}\Big|_{0}^{1} = \frac{1}{3} - \frac{1}{4} + \frac{1}{3} - \frac{1}{4} = \frac{1}{6}$$

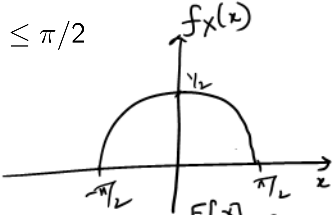
$$\int x^n\,dx = \frac{x^{n+1}}{n+1}$$

$$|x| = \begin{cases} -x, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases}$$

Problem

A continuous random variable X has PDF

$$f_X(x) = \begin{cases} \frac{1}{2}\cos x, & -\pi/2 \le x \le \pi/2 \\ 0, & \text{otherwise.} \end{cases}$$



Find the CDF of X, $E[X]$, Var(X).

$$F_X(x) = \begin{cases} 0, & x < -\pi/2 \\ ?, & -\pi/2 \le x \le \pi/2 \\ 1, & x > \pi/2 \end{cases}$$

$E[X] = 0$

$$\int \cos x \, dx = \sin x$$

$\underline{-\pi/2 \le x \le \pi/2}$

$$F_X(x) = \int_{-\pi/2}^{x} f_X(u)\,du = \frac{1}{2} \sin x \Big|_{-\pi/2}^{x} = \frac{1}{2}\Big(\sin x - \underbrace{\sin(-\pi/2)}_{-1}\Big) = \frac{1+\sin x}{2}$$

$\int x\cos x\,dx = \cos x + x\sin x$

$\int x^2\cos x\,dx = x^2\sin x + 2x\cos x - 2\sin x$

$$E[X] = \int_{-\pi/2}^{\pi/2} x\,\frac{1}{2}\cos x\,dx = \frac{1}{2}\left(\cos x + x\sin x\right)\Big|_{-\pi/2}^{\pi/2} = \frac{1}{2}\Big[\big(\overset{0}{\cancel{\cos\tfrac{\pi}{2}}} + \tfrac{\pi}{2}\overset{1}{\cancel{\sin\tfrac{\pi}{2}}}\big) - \big(\underset{0}{\cos(-\tfrac{\pi}{2})} - \tfrac{\pi}{2}\underbrace{\sin(-\tfrac{\pi}{2})}_{-1}\big)\Big]$$

$$= 0$$

$$\mathrm{Var}(X) = E[X^2] = \int_{-\pi/2}^{\pi/2} x^2\cdot\frac{1}{2}\cos x\,dx = \frac{1}{2}\left(x^2\sin x + 2x\cos x - 2\sin x\right)\Big|_{-\pi/2}^{\pi/2}$$

$$= \frac{1}{2}\left[\left(\frac{\pi^2}{4} - 2\right) - \left(-\frac{\pi^2}{4} + 2\right)\right]$$

$$= \frac{\pi^2}{4} - 2$$

Multiple discrete/continuous random variables

Andrew Thangaraj

IIT Madras

Subsection 1

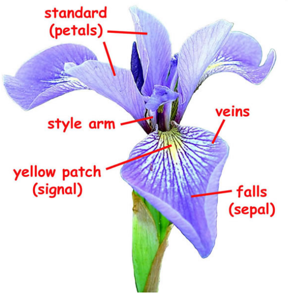
Motivation

Iris data set

- First used by R. A. Fisher
  - Wikipedia: https://en.wikipedia.org/wiki/Ronald_Fisher
    - "a genius who almost single-handedly created the foundations for modern statistical science"
    - "the single most important figure in 20th century statistics"

Iris data set

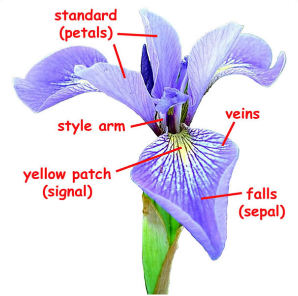
- First used by R. A. Fisher
  - Wikipedia: https://en.wikipedia.org/wiki/Ronald_Fisher
    - "a genius who almost single-handedly created the foundations for modern statistical science"
    - "the single most important figure in 20th century statistics"
- Iris flower
  - 3 classes of irises: 0, 1 and 2
    - 50 instances in each class
  - Data (cm)
    - sepal length (SL), sepal width (SW), petal length (PL), petal width (PW)
  - Classification
    - Given data, find class



(image source: fs.fed.us)

Iris data set

- First used by R. A. Fisher
  - Wikipedia: https://en.wikipedia.org/wiki/Ronald_Fisher
    - "a genius who almost single-handedly created the foundations for modern statistical science"
    - "the single most important figure in 20th century statistics"
- Iris flower
  - 3 classes of irises: 0, 1 and 2
    - 50 instances in each class
  - Data (cm)
    - sepal length (SL), sepal width (SW), petal length (PL), petal width (PW)
  - Classification
    - Given data, find class



(image source: fs.fed.us)

How to statistically describe (class, SL, SW, PL, PW)?

Iris data

Class 0

SL	SW	PL	PW
5.1	3.5	1.4	0.2
4.9	3.0	1.4	0.2
4.7	3.2	1.3	0.2
4.6	3.1	1.5	0.2
5.0	3.6	1.4	0.2
⋮	⋮	⋮	⋮

Class 1

SL	SW	PL	PW
7.0	3.2	4.7	1.4
6.4	3.2	4.5	1.5
6.9	3.1	4.9	1.5
5.5	2.3	4.0	1.3
6.5	2.8	4.6	1.5
⋮	⋮	⋮	⋮

Class 2

SL	SW	PL	PW
6.3	3.3	6.0	2.5
5.8	2.7	5.1	1.9
7.1	3.0	5.9	2.1
6.3	2.9	5.6	1.8
6.5	3.0	5.8	2.2
⋮	⋮	⋮	⋮

Iris data

Class 0

SL	SW	PL	PW
5.1	3.5	1.4	0.2
4.9	3.0	1.4	0.2
4.7	3.2	1.3	0.2
4.6	3.1	1.5	0.2
5.0	3.6	1.4	0.2
⋮	⋮	⋮	⋮

Class 1

SL	SW	PL	PW
7.0	3.2	4.7	1.4
6.4	3.2	4.5	1.5
6.9	3.1	4.9	1.5
5.5	2.3	4.0	1.3
6.5	2.8	4.6	1.5
⋮	⋮	⋮	⋮

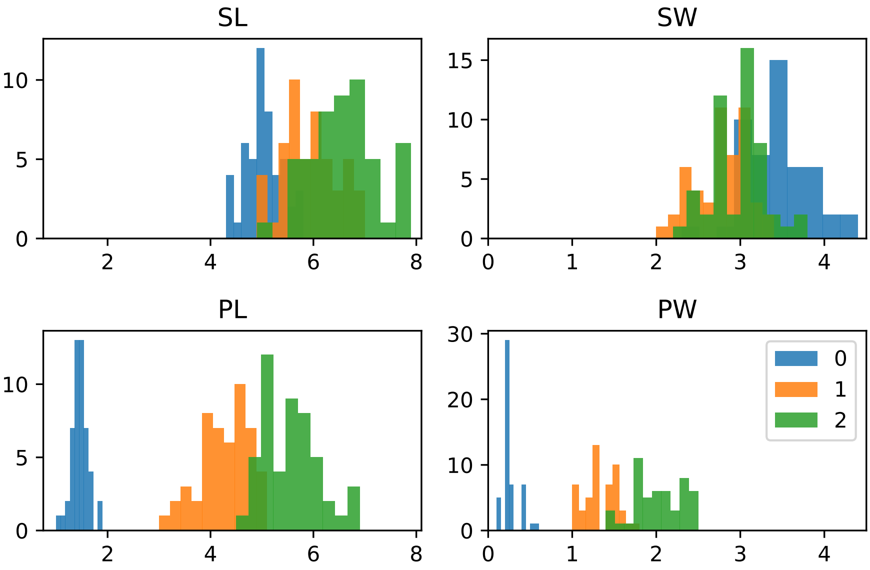
Class 2

SL	SW	PL	PW
6.3	3.3	6.0	2.5
5.8	2.7	5.1	1.9
7.1	3.0	5.9	2.1
6.3	2.9	5.6	1.8
6.5	3.0	5.8	2.2
⋮	⋮	⋮	⋮

Summary: min-max, avg, stdev

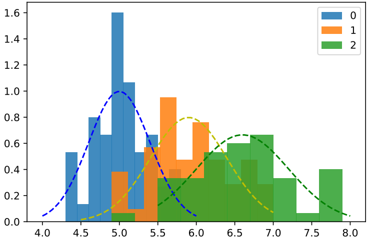
	SL summary	SW summary	PL summary	PW summary
0	4.3-5.8,5.0,0.4	2.3-4.4,3.4,0.4	1.0-1.9,1.5,0.2	0.1-0.6,0.3,0.1
1	4.9-7.0,5.9,0.5	2.0-3.4,2.8,0.3	3.0-5.1,4.3,0.5	1.0-1.8,1.3,0.2
2	4.9-7.9,6.6,0.6	2.2-3.8,3.0,0.3	4.5-6.9,5.6,0.6	1.4-2.5,2.0,0.3

Histograms



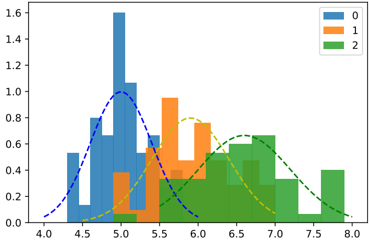
SL
10
5
0
2
4
6
8
SW
15
10
5
0
0
1
2
3
4
PL
10
5
0
2
4
6
8
PW
30
20
10
0
0
1
2
3
4
0
1
2

How to model sepal length and class of iris?



- density histograms of sepal length for three classes
- continuous approximations shown as dotted lines

How to model sepal length and class of iris?



- density histograms of sepal length for three classes
- continuous approximations shown as dotted lines

- Clearly, both are jointly distributed
- Class: discrete $\in \{0, 1, 2\}$
- Sepal length: continuous
  - distribution depends on class

Subsection 2

Joint distributions: Discrete and Continuous

Describing discrete-continuous joint distributions

- (X, Y): jointly distributed

Describing discrete-continuous joint distributions

- (X, Y): jointly distributed
- X: discrete with range T_X and PMF $p_X(x)$

Describing discrete-continuous joint distributions

- (X, Y): jointly distributed
- X: discrete with range T_X and PMF $p_X(x)$
- For each $x \in T_X$, we have a continuous random variable Y_x with density $f_{Y_x}(y)$

Describing discrete-continuous joint distributions

- (X, Y): jointly distributed
- X: discrete with range T_X and PMF $p_X(x)$
- For each $x \in T_X$, we have a continuous random variable Y_x with density $f_{Y_x}(y)$
- Y_x: Y given $X = x$, denoted $(Y|X = x)$

Describing discrete-continuous joint distributions

- (X, Y): jointly distributed
- X: discrete with range T_X and PMF $p_X(x)$
- For each $x \in T_X$, we have a continuous random variable Y_x with density $f_{Y_x}(y)$
- Y_x: Y given $X = x$, denoted $(Y|X = x)$
- $f_{Y_x}(y)$: conditional density of Y given $X = x$, denoted $f_{Y|X=x}(y)$

Describing discrete-continuous joint distributions

- (X, Y): jointly distributed
- X: discrete with range T_X and PMF $p_X(x)$
- For each $x \in T_X$, we have a continuous random variable Y_x with density $f_{Y_x}(y)$
- Y_x: Y given $X = x$, denoted $(Y|X = x)$
- $f_{Y_x}(y)$: conditional density of Y given $X = x$, denoted $f_{Y|X=x}(y)$
- Marginal density of Y

$$f_Y(y) = \sum_{x\in T_X} \overset{P(X=x)}{p_X(x)}\, \overbrace{f_{Y|X=x}(y)}^{\text{density}\,(Y|X=x)}$$

Problem

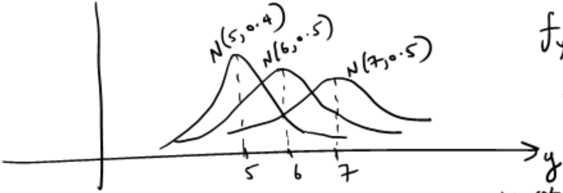
discrete uniform

conditional densities

Let $X \sim \text{Uniform}\{0, 1, 2\}$. Let $Y|X{=}0 \sim \text{Normal}(\underset{\mu=}{5}, \overset{\sigma=}{0.4})$, $Y|X{=}1 \sim \text{Normal}(\underset{\mu=}{6}, \overset{\sigma=}{0.5})$ and $Y|X{=}2 \sim \text{Normal}(\underset{\mu=}{7}, \overset{\sigma=}{0.6})$.

N ↔ Normal

- What is the marginal of Y?
- Suppose we observe Y to be around y_0. What can you say about X?



$$f_{Y|X=0}(y) = \frac{1}{\sqrt{2\pi} \times 0.4} e^{-\frac{(y-5)^2}{2(0.4)^2}}$$

$$f_{Y|X=1}(y) = \ ?$$

$$f_{Y|X=2}(y) = \ ?$$

$$f_Y(y) = \frac{1}{3} \cdot \frac{1}{\sqrt{2\pi} \cdot 0.4} e^{-\frac{(y-5)^2}{2(0.4)^2}} + \frac{1}{3} \cdot \frac{1}{\sqrt{2\pi} \cdot 0.5} e^{-\frac{(y-6)^2}{2(0.5)^2}} + \frac{1}{3} \cdot \frac{1}{\sqrt{2\pi} \cdot 0.6} e^{-\frac{(y-7)^2}{2(0.6)^2}}$$

"Not Gaussian"

called "Mixture of Gaussians"

Conditional probability of discrete given continuous

Definition

Suppose X and Y are jointly distributed with $X \in T_X$ being discrete with PMF $p_X(x)$ and conditional densities $f_{Y|X=x}(y)$ for $x \in T_X$. The conditional probability of X given $Y = y_0 \in \text{supp}(Y)$ is defined as

$$P(X{=}x|Y{=}y_0) = \frac{p_X(x) f_{Y|X=x}(y_0)}{\underbrace{f_Y(y_0)}_{= \sum_{x \in T_X} p_X(x) f_{Y|X=x}(y_0)}},$$

where f_Y is the marginal density of Y.

Conditional probability of discrete given continuous

Definition

Suppose X and Y are jointly distributed with $X \in T_X$ being discrete with PMF $p_X(x)$ and conditional densities $f_{Y|X=x}(y)$ for $x \in T_X$. The conditional probability of X given $Y = y_0 \in \text{supp}(Y)$ is defined as

$$P(X{=}x|Y{=}y_0) = \frac{p_X(x) f_{Y|X=x}(y_0)}{f_Y(y_0)},$$

where f_Y is the marginal density of Y.

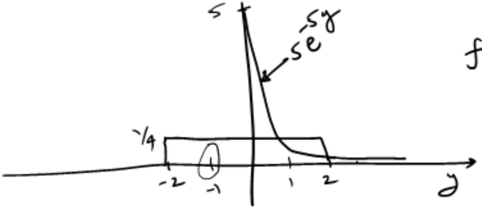
$P(A|B) \cdot P(B) = P(B|A) \cdot P(A) \rightarrow X = x$

- Similar to Bayes' rule: $P(X{=}x|Y{=}y_0) f_Y(y_0) = f_{Y|X=x}(y_0) p_X(x)$
- $X|Y{=}y_0$: "conditioned" discrete random variable
- When are X and Y independent? $f_{Y|X=x}$ is independent of x.
  - $f_Y = f_{Y|X=x}$ and $P(X{=}x|Y{=}y_0) = p_X(x)$

Problem

discrete — continuous

Let $X \sim \text{Uniform}\{-1, 1\}$. Let $Y|X = -1 \sim \text{Uniform}[-2, 2]$, $Y|X = 1 \sim \text{Exp}(5)$. Find the distribution of X given $Y = -1$, $Y = 1$, $Y = 3$.



$$f_Y(y) = \frac{1}{2} \cdot \underbrace{f_{Y|X=-1}(y)}_{\frac{1}{4},\ y\in(-2,2)} + \frac{1}{2} \underbrace{f_{Y|X=1}(y)}_{5e^{-5y},\ y>0}$$

$$= \begin{cases} 0, & y < -2 \\ \frac{1}{2}\cdot\frac{1}{4}, & -2 < y < 0 \\ \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2} 5e^{-5y}, & 0 < y < 2 \\ \frac{1}{2} 5e^{-5y}, & y > 2 \end{cases}$$

$X|Y=-1$: $P(X=-1|Y=-1) = \dfrac{p_X(-1)\cdot f_{Y|X=-1}(-1)}{f_Y(-1)} = \dfrac{\frac{1}{2}\cdot\frac{1}{4}}{\frac{1}{2}\cdot\frac{1}{4}} = 1$

$P(X=+1|Y=-1) = \dfrac{p_X(1) f_{Y|X=1}(-1)}{f_Y(-1)} = \dfrac{\frac{1}{2}\cdot 0}{\frac{1}{2}\cdot\frac{1}{4}} = 0$

$X|Y=1$: $P(X=-1|Y=1) = \dfrac{\frac{1}{2}\cdot\frac{1}{4}}{\frac{1}{2}\cdot\frac{1}{4}+\frac{1}{2}\cdot 5e^{-5}}$; $P(X=1|Y=1) = 1 - P(X=-1|Y=1)$

$= \dfrac{\frac{1}{2}\cdot 5e^{-5}}{\frac{1}{2}\cdot\frac{1}{4}+\frac{1}{2}\cdot 5e^{-5}}$

$X|Y=3$: $P(X=-1|Y=3) = 0$, $P(X=1|Y=3) = 1$

Problem

Suppose 60% of adults in the age group of 45-50 in a country are male and 40% are female. Suppose the height (in cm) of adult males in that age group in the country is Normal(160, $\sigma=$10), and that of females is Normal(150, $\sigma=$5). A random person is found to have a height of 155 cm. Is that person more likely to be male or female?

$$X \sim \{\overset{0.6}{M}, \overset{0.4}{F}\}$$

$$Y|X=M \sim N(160, \sigma=10) \quad f_{Y|X=M}(y) = \frac{1}{\sqrt{2\pi}\cdot 10} e^{-\frac{(y-160)^2}{2(10)^2}}$$

$$Y|X=F \sim N(150, \sigma=5) \quad f_{Y|X=F}(y) = \text{"}$$

$$f_Y(y) = ?$$

$$X|Y=155: \quad P(X=M|Y=155) = \frac{0.6 \times \frac{1}{\sqrt{2\pi}\cdot 10} e^{-\frac{5^2}{2\cdot 10^2}}}{0.6 \times \frac{1}{\sqrt{2\pi}\cdot 10}\cdot e^{-\frac{5^2}{2\cdot 10^2}} + 0.4 \cdot \frac{1}{\sqrt{2\pi}\cdot 5} e^{-\frac{5^2}{2\cdot 5^2}}} = ?$$

$$P(X=F|Y=155) = 1 - P(X=M|Y=155) = ?$$

Problem

Let $Y = X + Z$, where $X \sim \text{Uniform}\{-3, -1, 1, 3\}$ and $Z \sim \text{Normal}(0, \underset{\text{variance}}{\sigma^2})$ are independent. What is the distribution of Y? Find the distribution of $(X|Y = 0.5)$.

$f_{Y|X=-3}(y) = ?$

$Y|X=-3 \Leftrightarrow (-3+Z) \sim N(-3, \sigma^2)$

$Y|X=-1 \Leftrightarrow (-1+Z) \sim N(-1, \sigma^2)$

$Y|X=1 \Leftrightarrow N(1, \sigma^2)$; $Y|X=3 \Leftrightarrow N(3, \sigma^2)$

$\rightarrow$ rest is same as before.

Exercise: $X \sim \text{Unif}\{-3, -1, 1, 3\}$, $Z \sim N(0, \sigma^2)$ indep

$Y = XZ$