Chapter 3. Continuous Distribution

Math 3215 Summer 2023

Georgia Institute of Technology

Section 1.
Random Variables of the Continuous Type

Continuous Random Variables

$S \xrightarrow{X}$ number .

outcomes

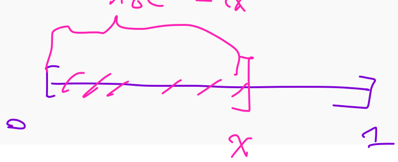
Let the random variable X denote the outcome when a point is selected at random from an interval $[0, 1]$.

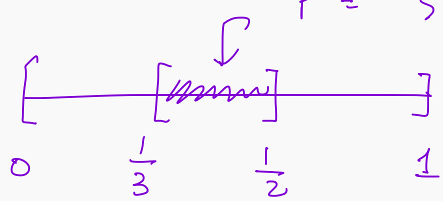
If the experiment is performed in a fair manner, it is reasonable to assume that the probability that the point is selected from an interval $[\frac{1}{3}, \frac{1}{2}]$ is

$P =$ size of $[\frac{1}{3}, \frac{1}{2}] = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$.

The cdf of X is

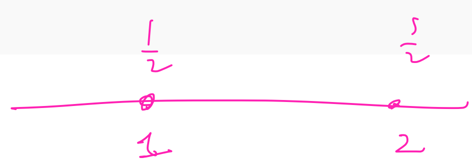
$\boxed{F(x) = x}$





$\mathbb{P}(X = \frac{1}{2}) = 0$

$\mathbb{P}(\frac{1}{6} \leq X \leq \frac{1}{3}) =$



$\frac{1}{2} = \mathbb{P}(X = 1)$, $\mathbb{P}(X = 2) = \frac{1}{2}$

Continuous Random Variables

Definition

We say a random variable X on a sample space S is a continuous random variable if there exists a function $f(x)$ such that

similar to pmf.

- $f(x) \geq 0$ for all x,
- $\int_{S(X)} f(x)\,dx = 1$, and
- For any interval $(a, b) \subset \mathbb{R}$,

$$\mathbb{P}(X \in (a,b)) = \mathbb{P}(a < X < b) = \int_a^b f(x)\,dx.$$

The function $f(x)$ is called the probability density function (pdf) of X.

density

finite, countable → discrete random variable.

S

otherwise

ex) $[0,1]$, $[0,\infty)$

$(-\infty,\infty) = \mathbb{R}$

$\boxed{\text{continuous random variable}}$.

X is conti.

there exists a density $f(x)$ need not to be conti.

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

Continuous Random Variables

The cdf of X is $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

The expectation (mean) of X is $E[X] = \int_{-\infty}^{\infty} x\,f(x)\,dx$

The variance of X is $Var(X) = E[(X - \underset{\mu = \mu_X}{E[X]})^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \sigma^2$

The standard deviation of X is $Std(X) = \sqrt{Var(X)} = \sigma = \sigma_X$

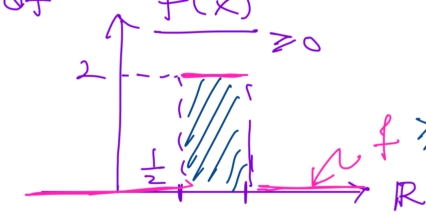
The moment generating function of X is

$$M(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)\,dx.$$

For pmf $f(k) \geq 0$, $\sum_k \overset{\mathbb{P}(X=k)}{f(k)} = 1 \Rightarrow f(k) \leq 1$

For pdf $f(x) \geq 0$, $\int_{-\infty}^{\infty} f(x)\,dx = 1$ but $f(x) > 1$ for some x

Ex



$\int f\,dx = 1$ but $\overset{\mathbb{P}(X=x)}{f(x)} > 1$ for $x \in [\frac{1}{2}, 1]$

Continuous Random Variables

Properties

The pmf of a discrete random variable is bounded by 1. But for pdf, $f(x)$ can be greater than 1.

For cdf F, we have $F'(x) = f(x)$ where F is differentiable at x.

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

$$\frac{d}{dx} F(x) = \frac{d}{dx} \int_{-\infty}^{x} f(t)\,dt = f(x)$$

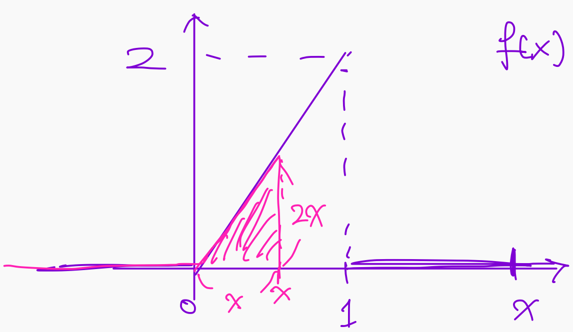
Fundamental Thm of Calculus.

Continuous Random Variables

Example

Let X be a continuous random variable with a pdf $\overset{f}{\cancel{g}}(x) = 2x$ for $0 < x < 1$.

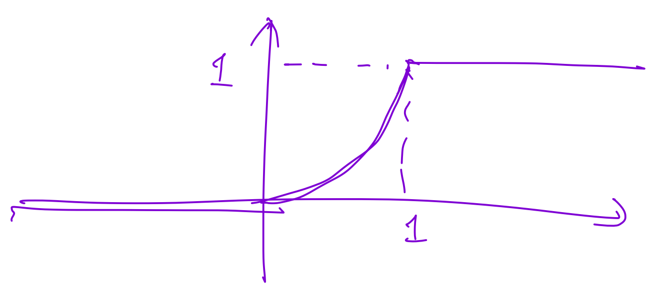
Find the cdf and the expectation.



$$f(x) = \begin{cases} 2x & \text{if } 0 < x < 1 \\ 0 & \text{o.w.} \end{cases}$$

$$F(x) = \underline{P}(X \le x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & , x \le 0 \\ \frac{1}{2} \cdot x \cdot (2x) = x^2 & , 0 < x < 1 \\ 1 & , x \ge 1 \end{cases}$$

$$F(x) = \begin{cases} 0 & x \le 0 \\ x^2 & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$$



Continuous Random Variables

Example

Let X be a continuous random variable with a pdf $\underbrace{\cancel{g}(x)}_{}$ f $= 2x$ for $0 < x < 1$.

Find the cdf and the expectation.

$$E[X] = \int_{-\infty}^{\infty} x \cdot \underline{\underline{f(x)}}\, dx = \int_{0}^{1} x \cdot 2x \, dx$$

$$= \int_{0}^{1} 2\boxed{x^2}\, dx = \left[2 \cdot \frac{1}{3} \cdot x^3\right]_0^1 = \frac{2}{3}.$$

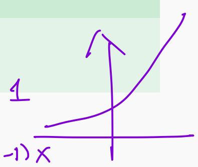
Continuous Random Variables

Example

Let X have the pdf $f(x) = xe^{-x}$. Find the mgf.

$$f(x) = \begin{cases} xe^{-x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$t>1$



$\frac{1}{(t-1)} e^{(t-1)x}$

∞

$$M(t) = \mathbb{E}[e^{tX}] = \int_0^\infty e^{tx} f(x)\, dx = \int_0^\infty x \cdot e^{(t-1)x}\, dx$$

$x \to 1$ $\qquad$ $t<1$

$$\overset{\text{IBP}}{=} \underbrace{\left[x \cdot \frac{1}{(t-1)} e^{(t-1)x} \right]_0^\infty}_{0} - \int_0^\infty 1 \cdot \frac{1}{(t-1)} e^{(t-1)x}\, dx$$

$$\int u(x) \cdot v'(x)\, dx = u(x) V(x) - \int u'(x) \cdot V(x)\, dx$$

$$= \left[-\frac{1}{(t-1)^2} e^{\overset{<0}{(t-1)}x} \right]_0^\infty = \frac{1}{(t-1)^2}, \quad t<1.$$

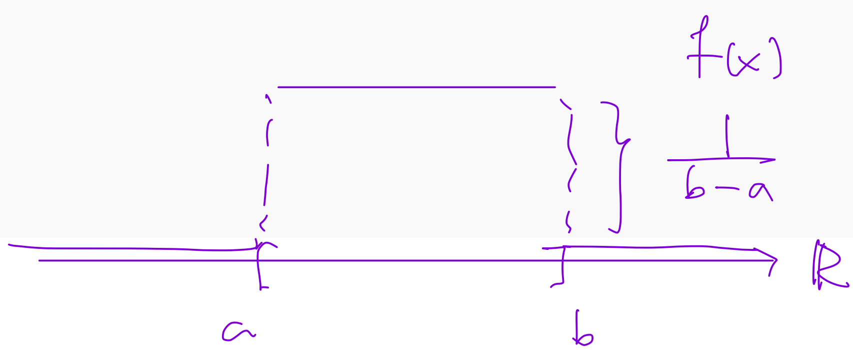
Uniform Random Variables

Definition

X is a uniform random variable if its pdf is constant on its support.

If its support is $[a, b]$, then the pdf is

We denote by $X \sim U(a, b)$.



Uniform Random Variables

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{o.w.} \end{cases}$$

Theorem

If $X \sim U(a, b)$, then

$\mathbb{E}[X] = \frac{a+b}{2}$

$\mathrm{Var}[X] = \frac{1}{12}(a-b)^2$

$M(t) =$ Exercise.

$$\mathbb{E}[X] = \int_a^b \frac{1}{b-a} \cdot x \, dx = \frac{1}{b-a} \cdot \left[\frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2 \cdot (b-a)}$$

$$= \frac{a+b}{2}$$

$b^2 - a^2 = (b-a) \cdot (b+a)$

$$\mathbb{E}[X^2] = \int_a^b \frac{1}{b-a} x^2 \, dx = \frac{1}{(\cancel{b-a})} \cdot \frac{1}{3} \cdot \underbrace{(b^3 - a^3)}_{(\cancel{b-a}) \cdot (a^2+ab+b^2)}$$

$$= \frac{1}{3}(a^2 + ab + b^2)$$

$$\mathrm{Var}(X) = \frac{1}{3}(a^2 + ab + b^2) - \frac{1}{4}(a^2 + 2ab + b^2)$$

$$= \frac{1}{12}(a^2 - 2ab + b^2) = \frac{(a-b)^2}{12}.$$

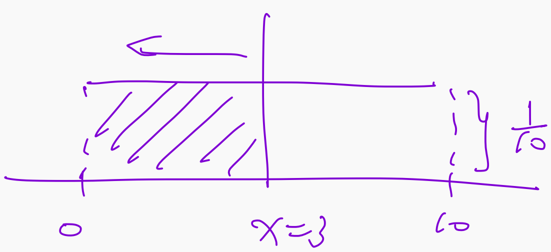
Uniform Random Variables

$X \sim \text{Unif}(0, 10)$

Example

If X is uniformly distributed over $(0, 10)$, calculate $\mathbb{P}(X < 3)$, $\mathbb{P}(X > 6)$, and $\mathbb{P}(3 < X < 8)$.

$\mathbb{P}(X < 3) = 3 \cdot \frac{1}{10} = 0.3$

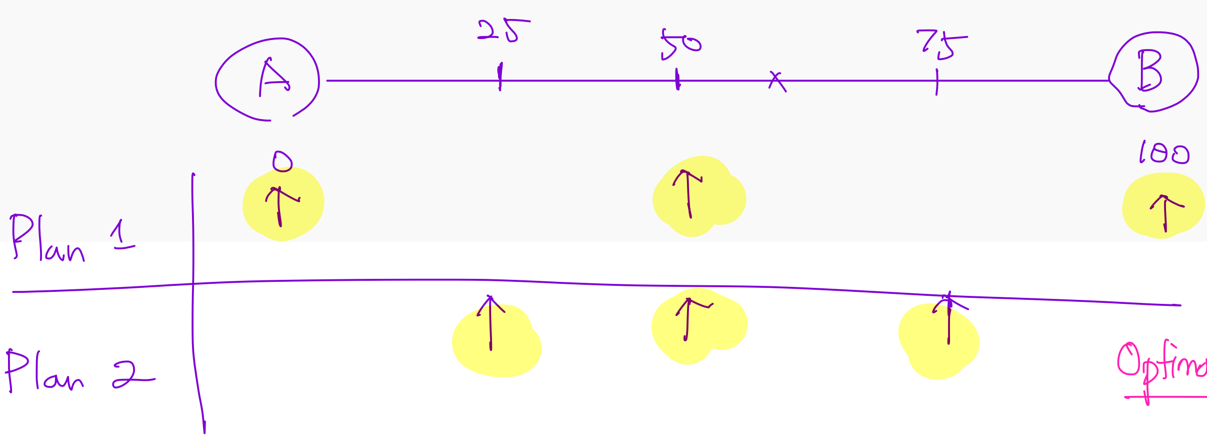


$\mathbb{P}(X > 6) = \frac{4}{10}$, $\mathbb{P}(3 < X < 8) = \frac{5}{10}$.

Uniform Random Variables

Example

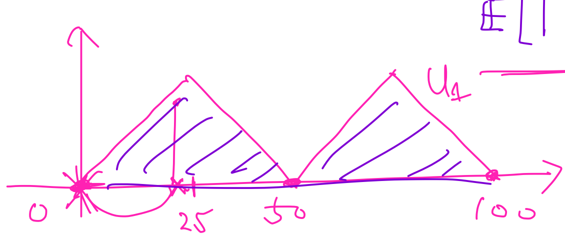
A bus travels between the two cities A and B, which are 100 miles apart. If the bus has a breakdown, the distance from the breakdown to city A has a $U(0, 100)$ distribution. There are bus service stations in city A, in B, and in the center of the route between A and B. It is suggested that it would be more efficient to have the three stations located 25, 50, and 75 miles, respectively, from A. Do you agree? Why?

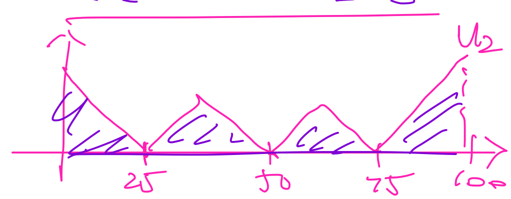


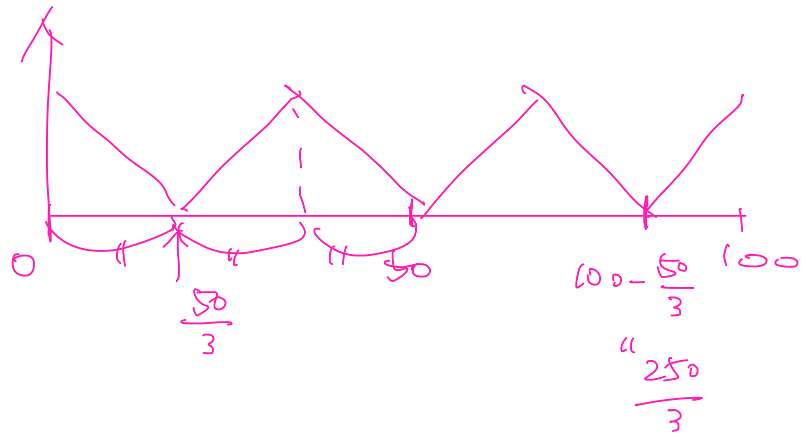
Optimal Transport.

$X \sim U(0, 100)$

$u_1(X) = E[|X - S_1|] \quad > \quad u_2(X) = E[|X - S_2|]$







Percentile

Ex

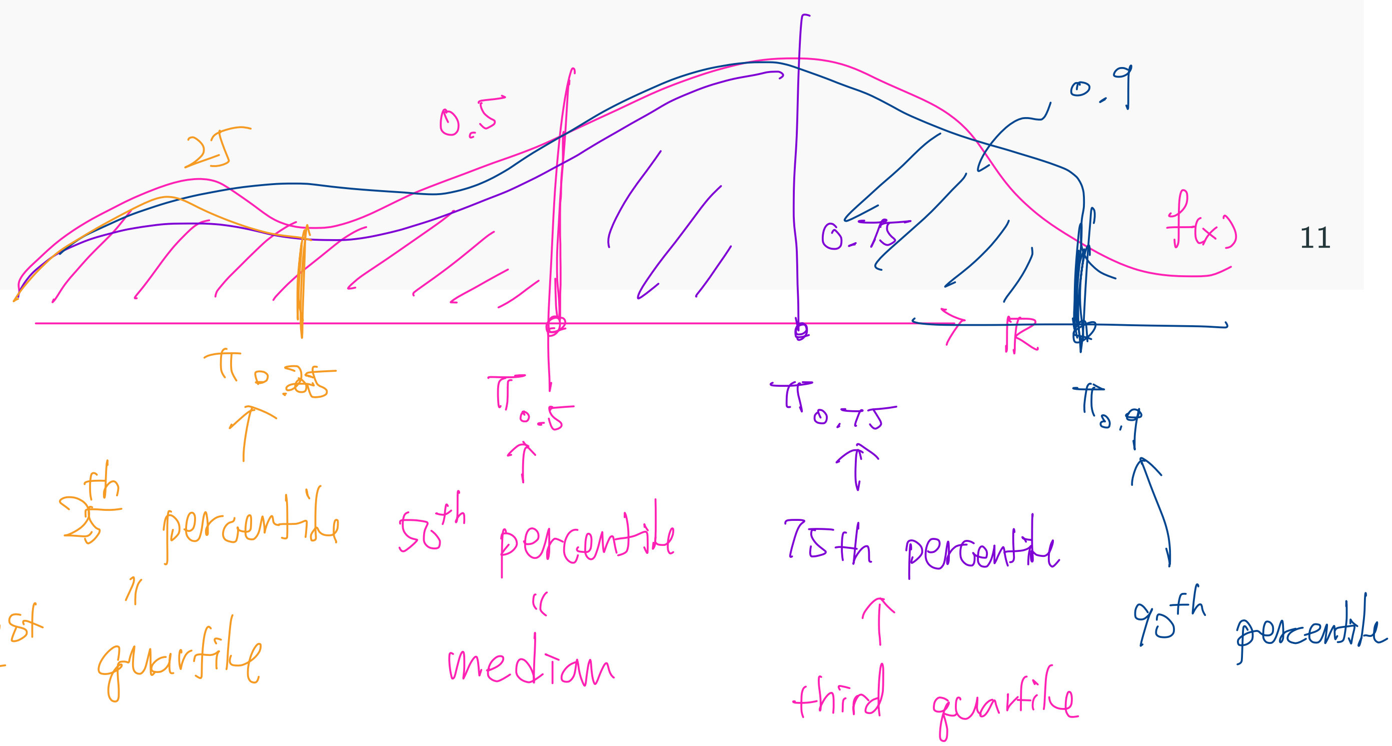
$= 100 \cdot 0.75$

75th percentile is $\pi_{0.75}$ s.t. $F(\pi_{0.75}) = 0.75$ (cdf)

The ($100p$)-th percentile is a number π_p such that $F(\pi_p) = p$.

For example, the 50th percentile is the number $\pi_{\frac{1}{2}} = q_2$ such that $F(\pi_{\frac{1}{2}}) = \frac{1}{2}$ and this is called the median.

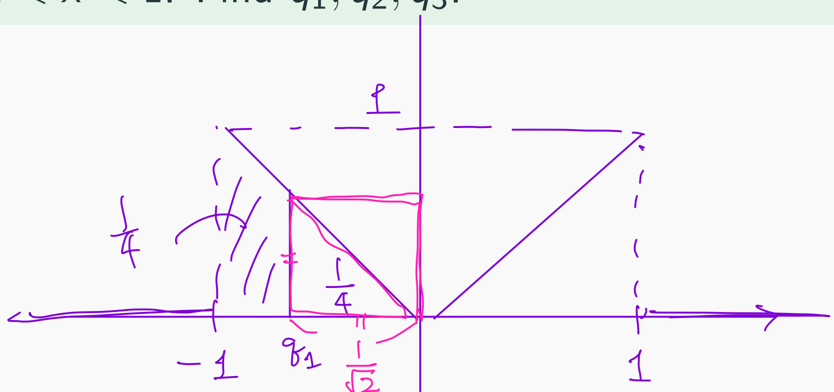
The 25th and 75th percentiles are called the first and third quartiles, respectively, and are denoted by $q_1 = \pi_{0.25}$ and $q_3 = \pi_{0.75}$.



Percentile

Example

Let X be a continuous random variable with pdf $f(x) = |x|$ for $-1 < x < 1$. Find q_1, q_2, q_3.

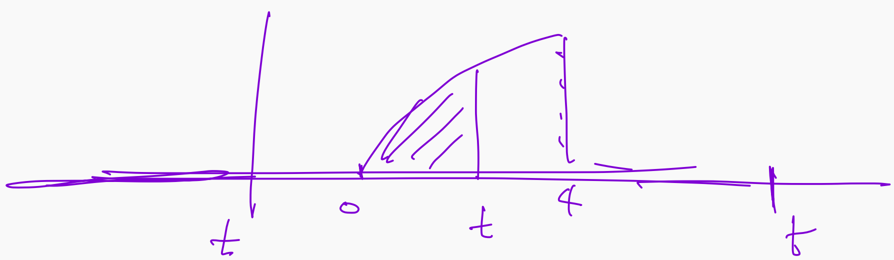


$q_2 = 50^{th}$ percentile $=$ median $= \pi_{0.5} = 0$

$q_1 = 25^{th}$ percentile $= 1^{st}$ quartile $= \pi_{0.25} = -\frac{1}{\sqrt{2}}$

$q_3 = \frac{1}{\sqrt{2}}$

Exercise



Let $f(x) = c\sqrt{x}$ for $0 \leq x \leq 4$ be the pdf of a random variable X.

Find c, the cdf of X, and $\mathbb{E}[X]$.

$$1 = \int_0^4 c\underbrace{\sqrt{x}}_{= x^{\frac{1}{2}}}\,dx = c\cdot\left[\frac{2}{3}\cdot x^{\frac{3}{2}}\right]_0^4 = c\cdot\frac{2}{3}\cdot 8 \quad \therefore\ c = \frac{3}{16}.$$

$$F(t) = \mathbb{P}(X \leq t) = \int_0^t c\cdot\sqrt{x}\,dx = \left[c\cdot\frac{2}{3}\cdot x^{\frac{3}{2}}\right]_0^t$$

$\underline{0 \leq t \leq 4}$

$$= \frac{\cancel{3}}{{}_8\cancel{16}}\cdot\frac{\cancel{2}}{\cancel{3}}\cdot t^{\frac{3}{2}}$$

$$= \frac{1}{8}\,t^{\frac{3}{2}}$$

$$F(t) = \begin{cases} 0 & t < 0 \\ \frac{1}{8}\cdot t^{\frac{3}{2}} & 0 \leq t \leq 4 \\ 1 & t \geq 4 \end{cases}$$

Section 2.
The Exponential, Gamma, and Chi-Square Distributions

exponential random variable

W 1^{st} occurrence

\# of customer in $[0,1] \sim$ Poisson (λ)

Poisson (λt)

$1 - \mathbb{P}(W > t)$

$W \sim Exp(\lambda)$ PDF? CDF first. $F(t) = \mathbb{P}(W \leq t)$

Exponential random variables

Consider a Poisson random variable X with parameter λ.

This represents the number of occurrances in a given interval, say $[0, 1]$.

If $\lambda = 5$, that means the expected number of occurrances in $[0, 1]$ is 5.

Let W be the waiting time for the first occurrence. Then,

$$\mathbb{P}(W > t) = \mathbb{P}(\text{no occurrences in } [0, t]) = \mathbb{P}(Y = 0)$$

for $t > 0$.

$$= e^{-\lambda t} \cdot \frac{(\lambda t)^0}{0!}$$

$$= e^{-\lambda t}$$

Y = \# of events in $[0,t] \sim$ Poisson (λt)

$F(t) = 1 - e^{-\lambda t}$ PDF $= f(t) = \lambda e^{-\lambda t}$, $t \geq 0$.

$F(t) = \int_{-\infty}^{t} f(s)\, ds$ PDF

$F'(t) = f(t)$

W = waiting time of 1^{st} occurrence $\sim Exp(\lambda)$

0 W 1

Exponential random variables

Definition

We say X is an exponential random variable with parameter λ (or mean θ where $\lambda = \frac{1}{\theta}$) if its pdf is

$$f(x) = \lambda e^{-\lambda x}$$

for $x \geq 0$ and otherwise 0. Here, λ is the parameter and θ is the mean.

$f(t) = \lambda e^{-\lambda t}, \quad t \geq 0 \qquad X \sim Exp(\lambda)$

$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx = \int_0^{\infty} x \lambda e^{-\lambda x}\, dx, \qquad \lambda x = t, \quad \lambda\, dx = dt, \quad dx = \frac{1}{\lambda} dt$

$= \frac{1}{\lambda} \int_0^{\infty} \underbrace{t}_{\downarrow\, 1} \cdot \underbrace{e^{-t}}_{\uparrow\, -e^{-t}}\, dt \underset{\text{IBP}}{=} \frac{1}{\lambda}\left(\underbrace{\left[-t e^{-t}\right]_0^{\infty}}_{=0} + \underbrace{\int_0^{\infty} e^{-t}\, dt}_{= \left[-e^{-t}\right]_0^{\infty} = 1} \right)$

$= \frac{1}{\lambda} = \theta.$

Exponential random variables

Theorem

Suppose that X is an exponential random variable with parameter $\lambda = \frac{1}{\theta}$.

$\mathbb{E}[X] = \frac{1}{\lambda} = \theta$

$\text{Var}[X] = \frac{1}{\lambda^2} = \theta^2 \quad = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

$M(t) = \frac{\lambda}{\lambda - t} = \frac{1}{1 - \theta t}$

$M(t) = \mathbb{E}[e^{tX}] = \int_0^{\infty} e^{tx} \cdot \lambda e^{-\lambda x}\, dx$

$= \lambda \int_0^{\infty} e^{-\underbrace{(\lambda - t)}_{>0} x}\, dx$

$\underset{\text{if } \lambda - t > 0}{=} \frac{\lambda}{\lambda - t} \underset{\theta = \frac{1}{\lambda}}{=} \frac{1}{1 - \theta t}.$

Exponential random variables

$f(x) = \lambda e^{-\lambda x} = \frac{1}{20} \cdot e^{-\frac{x}{20}}$

Example

Let X have an exponential distribution with a mean $\underline{\theta = 20}$. $\lambda = \frac{1}{\theta} = \frac{1}{20}$

Find $\mathbb{P}(X < 18)$.

$$\mathbb{P}(X < 18) = \int_{-\infty}^{18} f(x)\, dx = \int_0^{18} \frac{1}{20} e^{-\frac{x}{20}}\, dx$$

$$\mathbb{P}(X < 18) = F(18)$$

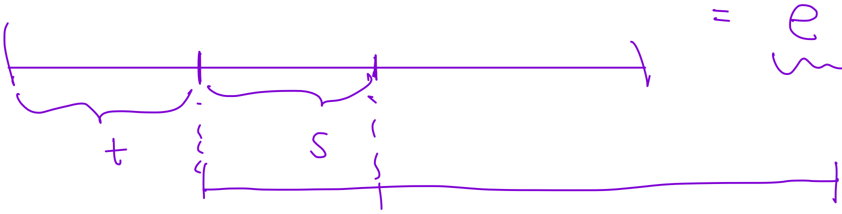
$$= \left[-e^{-\frac{x}{20}}\right]_0^{18} = 1 - e^{-\frac{18}{20}}.$$

① $F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X < x) = \underbrace{1 - e^{-\lambda x}}$

② $\mathbb{P}(X > x) = e^{-\lambda x}$

$$\mathbb{P}(X > t+s \mid X > t) = \frac{\mathbb{P}(X > t+s)}{\mathbb{P}(X > t)} = \frac{e^{-\lambda(\cancel{t}+s)}}{\cancel{e^{-\lambda t}}}$$

$$= e^{-\lambda s} = \mathbb{P}(X > s)$$



memoryless property.

$$\mathbb{P}(W>5) = e^{-\frac{1}{3}\cdot 5} = e^{-\frac{5}{3}}$$

Exponential random variables

$\lambda = 20$ Exp. # 20

0   $\frac{1}{1}$ = 1 hour   hour.

0   60   min $\lambda = \frac{1}{3}$ 0   1

Example

Customers arrive in a certain shop according to an approximate Poison process at a mean rate of 20 per hour.

What is the probability that the shopkeeper will have to wait more than five minutes for the arrival of the first customer?

X = # of customer in 1 hour $\sim$ Poisson(λ), $\lambda = 20$

$\Downarrow$

W = waiting time $\sim$ Exp(20)

$$\mathbb{P}\left(W > \frac{1}{12}\right) = e^{-20\cdot\frac{1}{12}} = e^{-\frac{5}{3}}.$$

$$\mathbb{P}(W>t) = e^{-\lambda t}$$

Binomial = # of success in n trials | Poisson = # of customers in [0,1]

Geometric = # of trials until 1st success | Exp. = Waiting time until 1st customer

Neg. Bin. = # of trials until r^{th} success | Gamma = Waiting time until r^{th} customers

Gamma random variables

Consider a Poisson random variable X with λ.

$\alpha = 1, 2, 3, \dots$

Let W be the waiting time until α-th occurrences, then its cdf is

$$F(t) = \mathbb{P}(W \le t) = 1 - \mathbb{P}(W > t) = 1 - \sum_{k=0}^{\alpha-1} \frac{(\lambda t)^k e^{-\lambda t}}{k!}.$$

differentiate in t

Thus, the pdf is

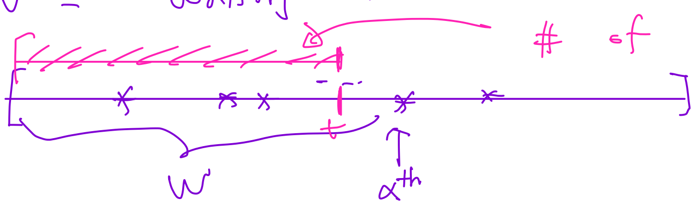
$$f(x) = \frac{\lambda(\lambda x)^{\alpha-1}}{(\alpha-1)!} e^{-\lambda x}.$$

This random variable is called a gamma random variable with λ and α where $\lambda = \frac{1}{\theta} > 0$.

This can be extended to non-integer $\alpha > 0$.

W = waiting time until α^{th} customers

# of customers in $[0,t] \sim$ Poisson(λt) $= Y$



$$\mathbb{P}(W > t) = \mathbb{P}(Y \le \alpha - 1) = \sum_{k=0}^{\alpha-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

$\alpha = 3$.
$$F(t) = 1 - e^{-\lambda t} - \lambda t e^{-\lambda t} - \frac{1}{2}(\lambda t)^2 e^{-\lambda t}$$

$$f(t) = F'(t) = \cancel{\lambda e^{-\lambda t}} - \cancel{\lambda e^{-\lambda t}} + \cancel{\lambda^2 t e^{-\lambda t}} - \cancel{\lambda^2 t e^{-\lambda t}} + \frac{1}{2} \cdot \lambda^3 \cdot t^2 \cdot e^{-\lambda t}$$

Gamma functions

$\Gamma(1) = \int_0^\infty y^{1-1} e^{-y} \, dy$

$= \int_0^\infty e^{-y} \, dy$

$= \left[-e^{-y} \right]_0^\infty = 1.$

The gamma function is defined by

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} \, dy$$

$-e^{-y}$

$(t-1)\, y^{t-2}$

$\Gamma(2) = \int_0^\infty y \cdot e^{-y} \, dy$ IBP

for $t > 0$.

By integration by parts, we have

$= \underbrace{\left[-y^{t-1} \cdot e^{-y} \right]_0^\infty}_{=0} + \int_0^\infty (t-1)\, y^{t-2} e^{-y} \, dy$

$= (t-1) \cdot \int_0^\infty y^{(t-1)-1} e^{-y} \, dy$?

$= (t-1) \cdot \Gamma(t-1)$

Def $X \sim \text{Gamma}(\lambda, \underline{\underline{\alpha}})$ if

$$f(x) = \frac{\lambda \cdot (\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \cdot e^{-\lambda x} \quad \text{for} \quad x \geq 0$$

Gamma functions

$\Gamma(t) = \Gamma(t-1) \cdot (t-1)$ $\Rightarrow$ $\Gamma(n) = (n-1)!$

In particular, $\Gamma(1) = 1$

$\Gamma(2) = (2-1) \cdot \Gamma(2-1) = \Gamma(1) = 1$

$\Gamma(3) = (3-1) \cdot \Gamma(3-1) = 2 \cdot \Gamma(2) = 2$

$\Gamma(n) = (n-1) \cdot \Gamma(n-1) = (n-1) \cdot (n-2) \cdot \Gamma(n-2) = (n-1) \cdot (n-2) \cdots \cdot \underbrace{\Gamma(1)}_{1}$

$= (n-1)!$

for integers n.

$X \sim \text{Gamma}(\lambda, \alpha)$

$$f = \frac{\lambda \cdot (\lambda x)^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}, \quad x \geq 0.$$

Gamma random variables

$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy$ // $\underline{\Gamma(t) = (t-1)\Gamma(t-1)}$

Theorem

$\mathbb{E}[X] = \frac{\alpha}{\lambda}$ $\leftarrow$ When computing, use definition of Gamma functions

$\text{Var}[X] = \frac{\alpha}{\lambda^2}$

$M(t) = \frac{1}{(1-\theta t)^\alpha}$ for $t \leq \frac{1}{\theta}$. $\theta = \frac{1}{\lambda}$

Gamma random variables

$$f(x) = \frac{\lambda \cdot (\lambda x)^{2-1}}{\underbrace{\Gamma(2)}_{=1}} e^{-\lambda x} = \frac{1}{9} \cdot x\, e^{-\frac{x}{3}}$$

Example

Suppose the number of customers per hour arriving at a shop follows a Poisson random variable with mean 20.

λ = mean of Poisson.

That is, if a minute is our unit, then $\lambda = \frac{1}{3}$.

θ = mean of Exp.

What is the probability that the second customer arrives more than five minutes after the shop opens for the day?

W = waiting time for 2^{nd} customers $\sim \text{Gamma}\left(\frac{1}{3}, 2\right)$

$$P(W > 5) = \int_5^\infty \frac{1}{9} x e^{-\frac{x}{3}} dx = \frac{1}{\cancel{3}} \int_{\frac{5}{3}}^\infty y\, e^{-y} \cancel{3}\, dy$$

$\frac{x}{3} = y$, $dx = 3\,dy$

$$= \int_{\frac{5}{3}}^\infty y\, e^{-y} dy = \left[-y e^{-y}\right]_{\frac{5}{3}}^\infty + \int_{\frac{5}{3}}^\infty e^{-y} dy$$

$-e^{-y}$

1

$$= \frac{5}{3} e^{-\frac{5}{3}} + e^{-\frac{5}{3}} = \frac{8}{3} e^{-\frac{5}{3}}.$$

Chi-square distribution

$\lambda = \frac{1}{2}$

Let X have a gamma distribution with $\theta = 2$ and $\alpha = r/2$, where r is a positive integer.

The pdf of X is

$$f(x) = \frac{1}{\Gamma(\frac{r}{2})2^{\frac{r}{2}}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}}$$

$\leftarrow$ many application in stat.

for $x > 0$.

χ^2

We say that X has a chi-square distribution with r degrees of freedom and we use the notation $X \sim \chi^2(r)$.

Exercise

$\lambda = \frac{1}{\theta}$

Let X have an exponential distribution with mean θ.

Compute $\mathbb{P}(X > 15|X > 10)$ and $\mathbb{P}(X > 5)$.

① $\mathbb{P}(X > t) = e^{-\lambda t} \Longrightarrow \mathbb{P}(X > 5) = e^{-\lambda \cdot 5} = e^{-5/\theta}$

② $\mathbb{P}(X > t+s \mid X > t) = \mathbb{P}(X > s)$

$\mathbb{P}(X > 10+5 \mid X > 10) = \mathbb{P}(X > 5) = e^{-5/\theta}$.

Section 3.
The Normal Distribution

Central Limit Theorem.

Gaussian random variables

$X \sim Exp(\lambda)$ $\quad f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$

Definition

We say X is a Gaussian random variable or has a normal distribution if its pdf is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \qquad \forall x \in \mathbb{R}$$

Here μ is the mean and σ is the standard deviation. We use the notation $X \sim N(\underset{\text{mean}}{\mu}, \underset{\text{variance}}{\sigma^2})$.

If $\mu = 0, \quad \sigma^2 = 1.$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{|x|^2}{2}}$$

$X \sim N(0,1)$ the standard normal (Gaussian).

$X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{|x-\mu|^2}{2\sigma^2}}$$

Gaussian random variables

Theorem

$\int_{\mathbb{R}} f(x)\, dx = 1$

$\mathbb{E}[X] = \mu$

$\text{Var}[X] = \sigma^2$

$M(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$

For $\mu = 0$, $\sigma^2 = 1$,

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{|x|^2}{2}} dx = 1 \quad \text{(Gaussian Integral)}$$

$Z \sim N(0, 1)$

① $Z \sim N(0,1) \Rightarrow X = \sigma Z + \mu \sim N(\mu, \sigma^2)$

② $X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Standard normal distribution

$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{|x|^2}{2}}$

In particular, if $\mu = 0$ and $\sigma^2 = 1$, then $Z \sim N(0,1)$ is called the standard normal random variable.

Example

Let Z ~~is~~ be $N(0,1)$.

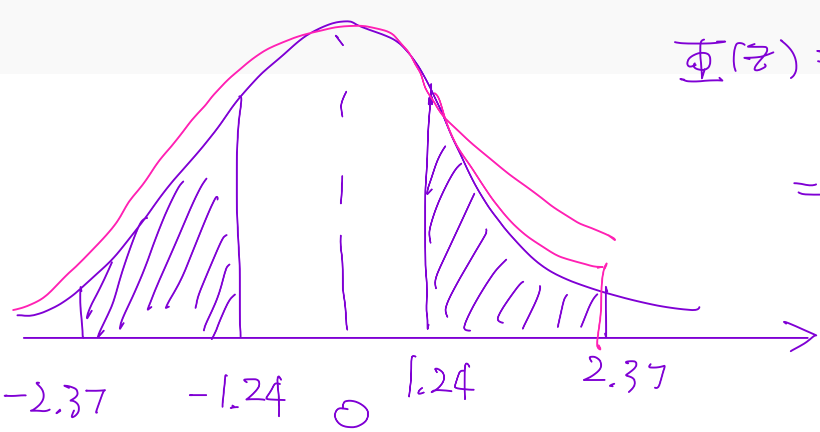
Find $\mathbb{P}(Z \leq 1.24)$, $\mathbb{P}(1.24 \leq Z \leq 2.37)$, and $\mathbb{P}(-2.37 \leq Z \leq -1.24)$.

$\mathbb{P}(Z \leq 1.24) =$ ____

$\mathbb{P}(-2.37 \leq Z \leq -1.24) = \Phi(2.37) - \Phi(1.24)$

$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{|x|^2}{2}} dx$

$= \mathbb{P}(Z \leq z)$



Standard normal distribution

Theorem

If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma}$ is the standard normal.

Standard normal distribution

$\leadsto X = \underline{4Z+3}$

$Z = \dfrac{X-3}{4} \sim N(0,1)$

Example

Let $X \sim N(3, 16)$. $\mu = 3$, $\sigma^2 = 16$, $\underline{\sigma = 4}$.

Find $\mathbb{P}(4 \leq X \leq 8)$, $\mathbb{P}(0 \leq X \leq 5)$, and $\mathbb{P}(-2 \leq X \leq 1)$.

$\mathbb{P}(4 \leq X \leq 8) = \mathbb{P}(\underline{4} \leq 4Z + \underline{3} \leq \underline{8})$

in terms of Z

$= \mathbb{P}(1 \leq 4Z \leq 5)$

$= \mathbb{P}(0.25 \leq Z \leq 1.25)$

$= \Phi(1.25) - \Phi(0.25)$

Standard normal distribution

Example

Let $X \sim N(25, 36)$. $\mu = 25$, $\sigma^2 = 36$, $\sigma = 6$

Find a constant c such that $\mathbb{P}(|X - 25| \le c) = 0.9544$. $c = 12$

$Z \sim N(0,1)$ $X = \sigma Z + \mu = \underbrace{6Z + 25}$

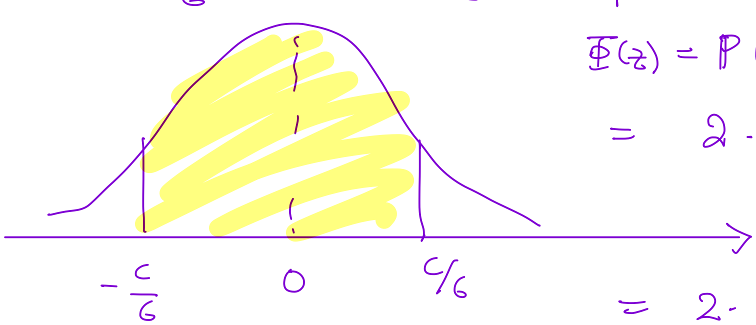
$\mathbb{P}(|X - 25| \le c)$

$= \mathbb{P}(|6Z| \le c) = \mathbb{P}(|Z| \le c/6)$

$= \mathbb{P}\left(-\frac{c}{6} \le Z \le \frac{c}{6}\right) = \Phi\left(\frac{c}{6}\right) - \Phi\left(-\frac{c}{6}\right)$

$\Phi(z) = \mathbb{P}(Z \le z)$

$= 2 \cdot \mathbb{P}\left(0 \le Z \le \frac{c}{6}\right)$



$\frac{c}{6} = 2$

$\Uparrow$

$= 2 \cdot \left(\mathbb{P}\left(Z \le \frac{c}{6}\right) - \mathbb{P}(Z \le 0)\right)$

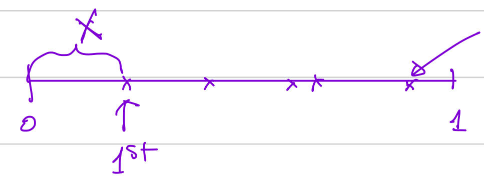
$\Phi\left(\frac{c}{6}\right) = \frac{1.9544}{2} = 0.9772$ $= 2 \cdot \left(\Phi\left(\frac{c}{6}\right) - \Phi(0)\right)$ $\Phi(0) \to 1/2$

$0.9544 = 2 \cdot \Phi\left(\frac{c}{6}\right) - 1$

Recall

- $X \sim \text{Exp}(\lambda)$ (rate), $\theta = \frac{1}{\lambda}$: the mean of X

\# of events $\sim$ Poisson (λ)



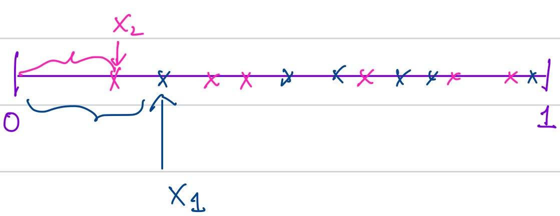
PMF : $f(t) = \lambda \cdot e^{-\lambda t}$, $t \geq 0$.

$E[X] = \frac{1}{\lambda} = \theta$

$Var(X) = \frac{1}{\lambda^2} = \theta^2$

- * $P(X > t+s \mid X > t) = P(X > s)$
- $P(X > t) = e^{-\lambda t}$
- * $X_1, X_2 \sim \text{Exp}(\lambda)$ Indep.

$Y = \min\{X_1, X_2\}$ = waiting time of first occurrence among two types of events.



: Exponential

$$P(Y > t) = P(\{X_1 > t\} \text{ and } \{X_2 > t\}) = P(X_1 > t) \cdot P(X_2 > t)$$

(and: intersection)

$\min\{X_1, X_2\}$

$$= e^{-\lambda t} \cdot e^{-\lambda t} = e^{-2\lambda t}$$

$Y \sim \text{Exp}(2\lambda)$.

Gamma = Waiting time until α^{th} event.

$$\text{PMF} = f(x) = \frac{\lambda (\lambda x)^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}$$

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx. \qquad (\alpha-1)!$$

$$\Gamma(\alpha) = (\alpha-1)\cdot\Gamma(\alpha-1).$$

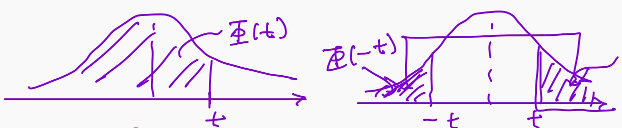
$X \sim N(\underset{\substack{\uparrow \\ \text{mean}}}{\mu}, \overset{\text{variance}}{\sigma^2})$ if $f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$-\infty < x < \infty$

If $\mu = 0$, $\sigma^2 = 1$, $Z \sim N(0,1)$: Standard normal.

If $X \sim N(\mu, \sigma^2)$ then $\begin{cases} X = \sigma Z + \mu \\ \frac{X-\mu}{\sigma} \sim N(0,1) \end{cases}$

Standard normal distribution



$\frac{d}{dt}\Phi(t) = f_Z(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$.

PDF of Z.

$\Phi(t) = P(Z \le t) =$ CDF of Z

$\Phi(-t) = 1 - \Phi(t)$

special case of Gamma(λ, α)

$\lambda = \frac{1}{2}$, $\theta = 2$, $\alpha = \frac{r}{2}$

r: degree of freedom.

Theorem

If Z is the standard normal, then Z^2 is $\chi^2(1)$.

$\Downarrow$

$X \sim \chi^2(r) = \text{Gamma}\left(\frac{1}{2}, \frac{r}{2}\right)$

$Z \sim N(0,1)$

$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$

$F_{Z^2}(t) = P(Z^2 \le t) = P(-\sqrt{t} \le Z \le \sqrt{t}) = \Phi(\sqrt{t}) - \Phi(-\sqrt{t})$

$(t \ge 0)$

$= \Phi(\sqrt{t}) - (1 - \Phi(\sqrt{t})) = 2 \cdot \Phi(\sqrt{t}) - 1$

$f_{Z^2}(t) = \frac{d}{dt} F_{Z^2}(t) = 2 \cdot \Phi'(\sqrt{t}) \cdot \frac{d}{dt}(\sqrt{t})$ (Chain rule)

PDF of Z^2.

$\Phi'(\sqrt{t}) = \frac{d}{ds}\Phi(s)\Big|_{s=\sqrt{t}}$

$= 2 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\sqrt{t})^2} \cdot \frac{1}{2\sqrt{t}} = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}t}$

$\sim \text{Gamma}\left(\underset{=\lambda}{\frac{1}{2}}, \underset{=\alpha}{\frac{1}{2}}\right)$

$\frac{1}{\sqrt{2x} \cdot \Gamma(\frac{1}{2})} e^{-\frac{x}{2}} = \frac{(\frac{1}{2}) \cdot (\frac{1}{2}x)^{\frac{1}{2}-1}}{\Gamma(\frac{1}{2})} e^{-\frac{1}{2}x}$

$\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Section 4.
Additional Models

of events   # of events $\Rightarrow$ Indep.   constant.

# of events = 0, 1 , $\mathbb{P}(\# = 1) \approx \lambda \cdot h$

$h \ll 1$

Weibull distribution

Recall the postulates of an approximate Poisson:

- The numbers of occurrences in nonoverlapping subintervals are independent.
- The probability of two or more occurrences in a sufficiently short subinterval is essentially zero.
- The probability of exactly one occurrence in a sufficiently short subinterval of length h is approximately λh.

Weibull distribution

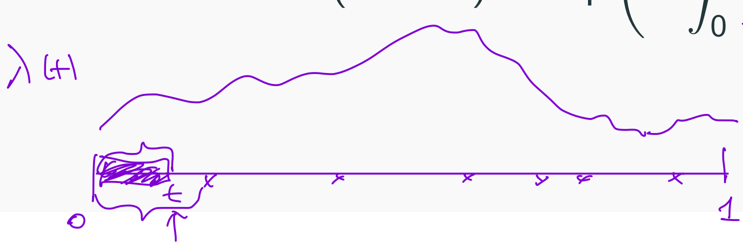
One can think the event occurrence as a failure and so λ can be understood as the failure rate.

Poisson distribution and its waiting time (exponential distribution) has a constant failure rate. λ

Sometimes, it is more natural to choose λ as a function of t in the last assumption.

Then the waiting time W for the first occurrence satisfies

$$\mathbb{P}(W > t) = \exp\left(-\int_0^t \lambda(w)\, dw\right).$$



$\leftarrow t$

$\int_0^t \underline{\lambda}\, ds$

$\parallel$

$$\mathbb{P}(W > t) = \mathbb{P}(\text{no occurrence in } [0,t]) \;\left(= e^{-\lambda t}.\right)$$

exp.

$$= e^{-\int_0^t \lambda(s)\, ds}$$

$$F_W(t) = \mathbb{P}(W \leq t) = 1 - \mathbb{P}(W > t) = 1 - e^{-\frac{t^\alpha}{\beta^\alpha}}$$

$$\Rightarrow \quad f_W(t) = \frac{d}{dt} F_W(t)$$

Weibull distribution

$$\mathbb{P}(W > t) = e^{-\int_0^t \frac{\alpha s^{\alpha-1}}{\beta^\alpha} ds} = e^{-\frac{1}{\beta^\alpha} \cdot t^\alpha}$$

Definition

If $\lambda(t) = \alpha \frac{t^{\alpha-1}}{\beta^\alpha}$, then the waiting time W for the first occurrence has the density

$$f_W(t) \; \cancel{g(t)} = \lambda(t) \exp\left(-\int_0^t \lambda(w)\, dw\right) = \underline{\alpha \frac{t^{\alpha-1}}{\beta^\alpha} \exp\left(-(\frac{t}{\beta})^\alpha\right)}.$$

W is called the Weibull random variable.

Weibull distribution

$$P(W>t) = e^{-\int_0^t 2s\,ds} = e^{-t^2}$$

Example

If $\lambda(t) = 2t^{1}$, then the waiting time W has the density

and it is a Weibull random variable with $\alpha = 2$ and $\beta = 1$.

If W_1, W_2 are independent Weibull with $\overset{=2}{\alpha}$ and $\overset{=1}{\beta}$ above, is the minimum of W_1, W_2 Weibull?

$$W \sim \text{Weibull}(\alpha, \beta), \qquad \lambda(t) = \frac{\alpha t^{\alpha-1}}{\beta^{\alpha}} = \frac{2t}{\beta^2} \quad (\alpha = 2)$$

$$\beta^2 = 1, \quad \beta = 1.$$

$$Z = \min\{W_1, W_2\}$$

$$F_Z(t) = P(Z \le t) = 1 - P(Z > t)$$

$$= 1 - P(\min\{W_1, W_2\} > t)$$

$$= 1 - \underbrace{P(W_1 > t)}\,\underbrace{P(W_2 > t)}$$

$$= 1 - e^{-t^2} \cdot e^{-t^2} = 1 - \underbrace{e^{-2t^2}}_{\shortparallel} \leftarrow \text{Weibull}$$

$$P(Z > t) = e^{-\int_0^t \lambda(s)\,ds}$$

$$\lambda(s) = 4s = \frac{2 \cdot s^{2-1}}{(1/2) = \beta^2} \qquad \alpha = 2$$

$$\therefore \beta = \frac{1}{\sqrt{2}}.$$

Weibull distribution

Theorem

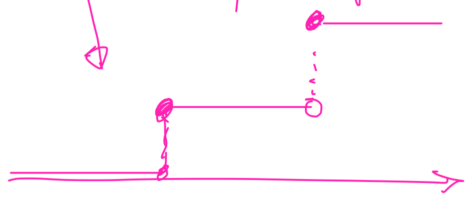
The mean of W is $\mu = \beta\Gamma(1 + \frac{1}{\alpha})$.

The variance is $\sigma^2 = \beta^2 \left(\Gamma(1 + \frac{2}{\alpha}) - \Gamma(1 + \frac{1}{\alpha})^2\right)$.

{ discrete RV — S is finite, or countable.
Conti. RV — there exists a density.

More RVs

only jump or flat



no jumps

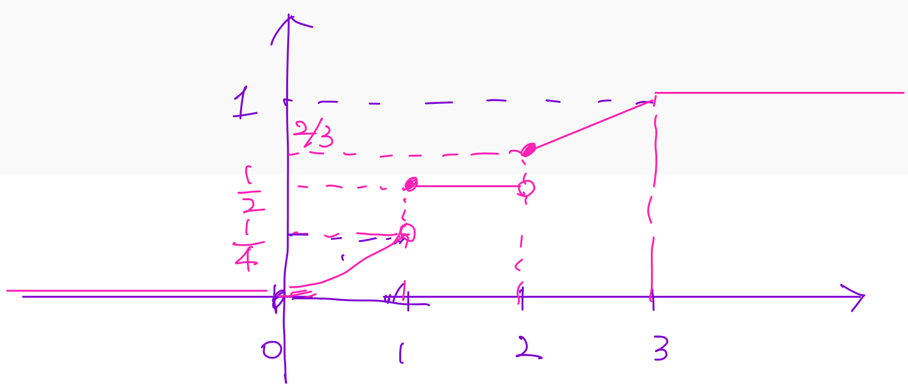
Mixed type random variables

Example

Suppose X has a cdf

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{4}, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ \frac{x}{3}, & 2 \leq x < 3 \\ 1, & x \geq 3. \end{cases}$$

Find $\mathbb{P}(0 < X < 1)$, $\mathbb{P}(0 < X \leq 1)$, and $\mathbb{P}(X = 1)$.



① $\mathbb{P}(0 < X < 1) = \frac{1}{4}$

② $\mathbb{P}(0 < X \leq 1) = \frac{1}{2} = \underbrace{\mathbb{P}(X \in (0,1))}_{\frac{1}{4}} + \underbrace{\mathbb{P}(X = 1)}_{\frac{1}{4}}$

③ $\mathbb{P}(X = 1) = \frac{1}{4}$.

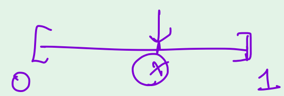
Mixed type random variables

Example

Consider the following game: A fair coin is tossed.

If the outcome is heads, the player receives \$2.

If the outcome is tails, the player spins a balanced spinner that has a scale from 0 to 1.



$U \sim \text{Unif}(0,1)$

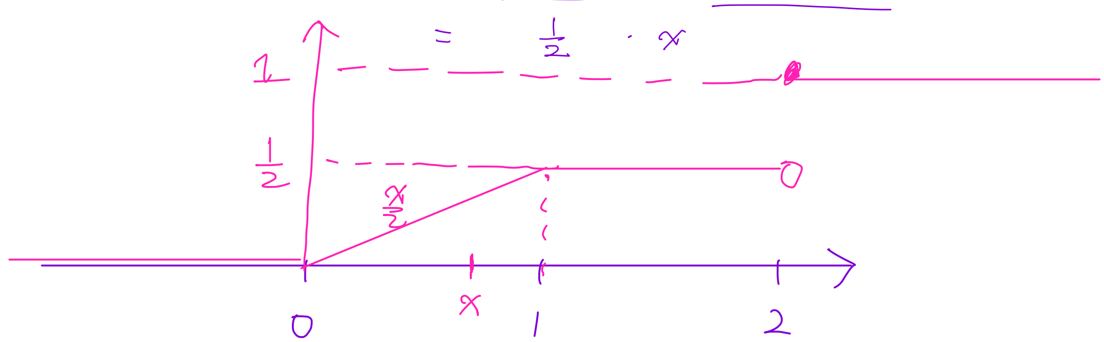
The player then receives that fraction of a dollar associated with the point selected by the spinner.

Let X be the amount received. Draw the graph of the cdf $F(x)$.

$$X = \begin{cases} 2, & \text{Heads} \\ U, & \text{Tails} \end{cases}$$

(if $0<x<1$)

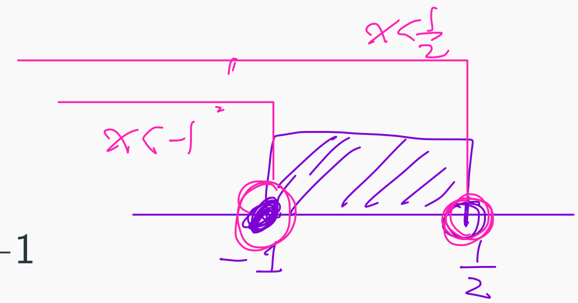
$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\text{Tails and } U \leq x)$$
$$= \underline{\mathbb{P}(\text{Tails})}\ \underline{\mathbb{P}(U \leq x)}$$
$$= \frac{1}{2} \cdot x$$



If $1 \le x < 2$ then $P(X \le x) = \frac{1}{2}$

$P(X \le 1.5) = P(\text{Tails}) = \frac{1}{2}$

Exercise



The cdf of X is given by

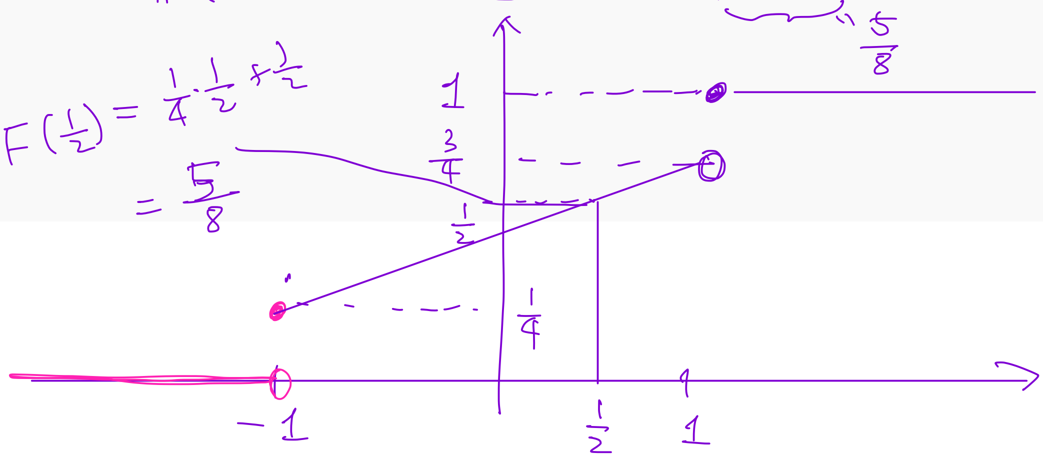
$$F(x) = \begin{cases} 0, & x < -1 \\ \frac{x}{4} + \frac{1}{2}, & -1 \le x < 1 \\ 1, & x \ge 1. \end{cases}$$

Find $\mathbb{P}(X < 0)$, $\mathbb{P}(X < -1)$, and $\mathbb{P}(-1 \le X < \frac{1}{2})$.

$P(-1 \le X < \frac{1}{2})$

$$\mathbb{P}(-1 \le X < \tfrac{1}{2}) = \underbrace{\mathbb{P}(X < \tfrac{1}{2})}_{\frac{5}{8}} - \underbrace{\mathbb{P}(X < -1)}_{0} = \frac{5}{8}.$$

$F(\frac{1}{2}) = \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{2} = \frac{5}{8}$



$\mathbb{P}(X < 0) = F(0) = \frac{1}{2}$

$\mathbb{P}(X < -1) = \underbrace{F(-1)}_{P(X \le -1)} - \mathbb{P}(X = -1) = 0$