Chapter 3. Continuous Distribution

Math 3215 Summer 2023

Georgia Institute of Technology

Section 1.
Random Variables of the Continuous Type

Continuous Random Variables

S outcomes $\xrightarrow{X}$ number.

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

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

size $= x$

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

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

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

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

S: finite, countable $\rightarrow$ discrete random variable.

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

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

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

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

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

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

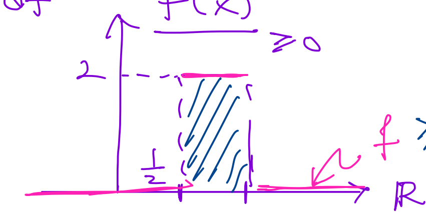
The moment generating function of X is

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

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

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 $\underbrace{f(x)}_{\mathbb{P}(X=x)} > 1$ 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.

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

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

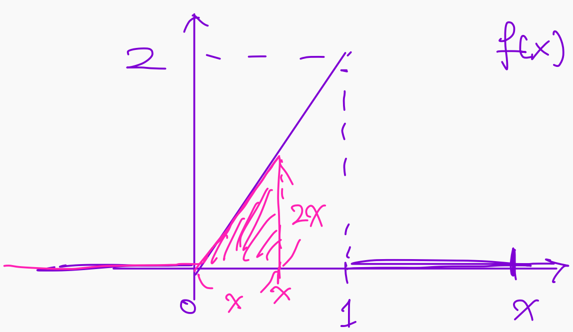
Fundamental Thm of Calculus.

Continuous Random Variables

Example

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

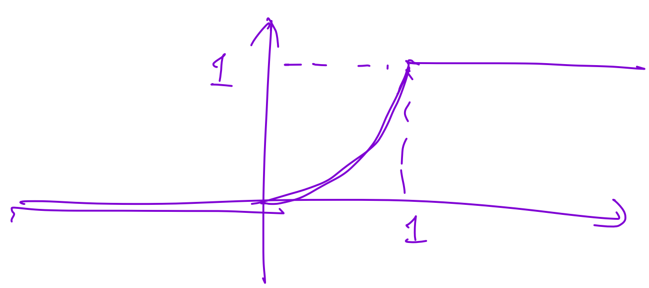
Find the cdf and the expectation.



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

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

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



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.

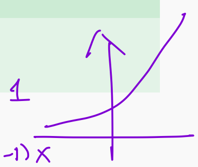
$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx = \int_0^1 x \cdot 2x\, dx$$

$$= \int_0^1 2\boxed{x^2}\, dx = \left[2 \cdot \frac{1}{3} \cdot x^3\right]_0^1 = \frac{2}{3}.$$

Example

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

$t>1$



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

$\frac{1}{(t-1)} e^{(t-1)x}$ ∞

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

1 $t<1$

$$\underset{\text{IBP}}{=} \Big[x \cdot \frac{1}{(t-1)} e^{(t-1)x} \Big]_0^\infty - \int_0^\infty 1 \cdot \frac{1}{(t-1)} e^{(t-1)x}\, dx$$

0

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

$V(x)$ $u'(x)$

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

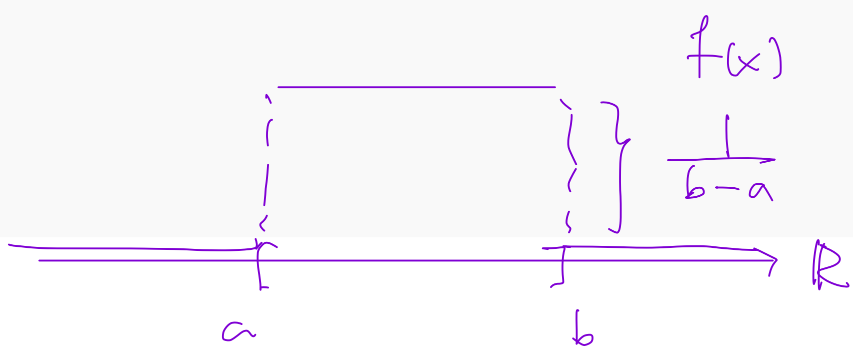
Uniform Random Variables

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

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

Theorem

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

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

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

$M(t) =$ Exercise.

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

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

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

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

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

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

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

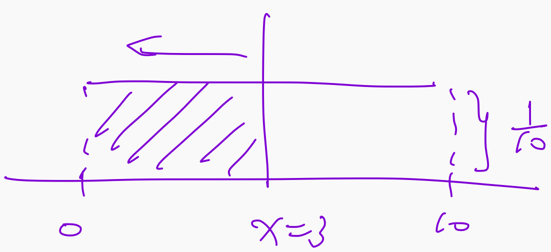
Uniform Random Variables

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

Example

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

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

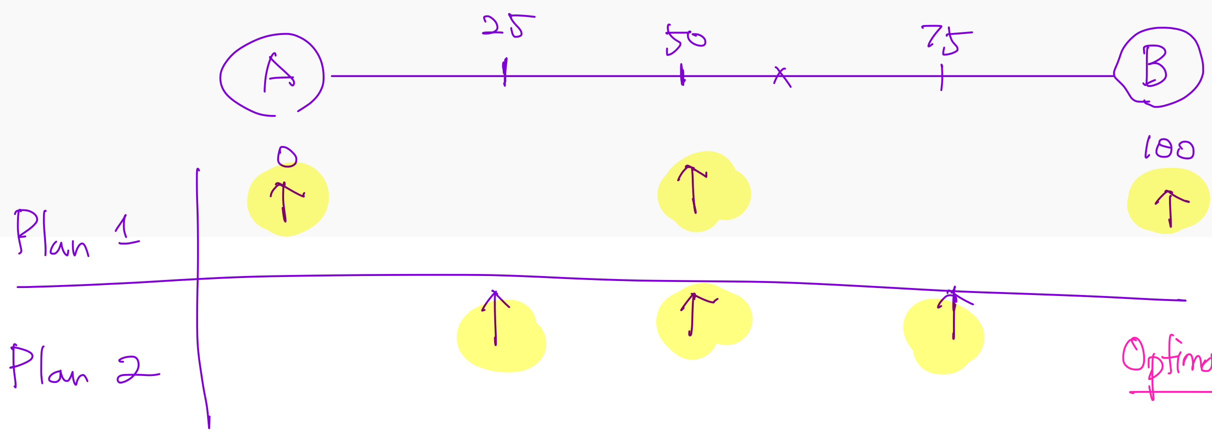


$\mathbb{P}(X > 6) = \frac{4}{10}$, $\mathbb{P}(3 < X < 8) = \frac{5}{10}$.

Uniform Random Variables

Example

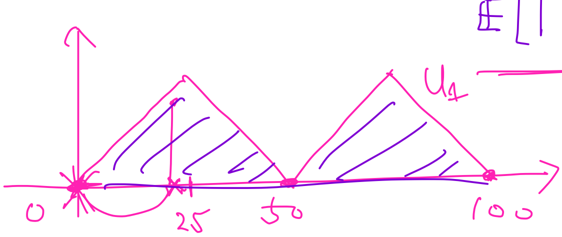
A bus travels between the two cities A and B, which are 100 miles apart. If the bus has a breakdown, the distance from the breakdown to city A has a $U(0, 100)$ distribution. There are bus service stations in city A, in B, and in the center of the route between A and B. It is suggested that it would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

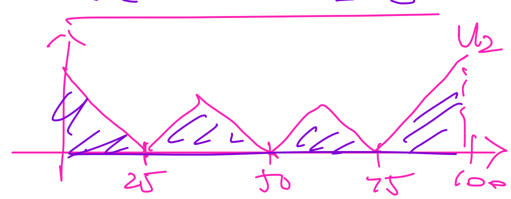


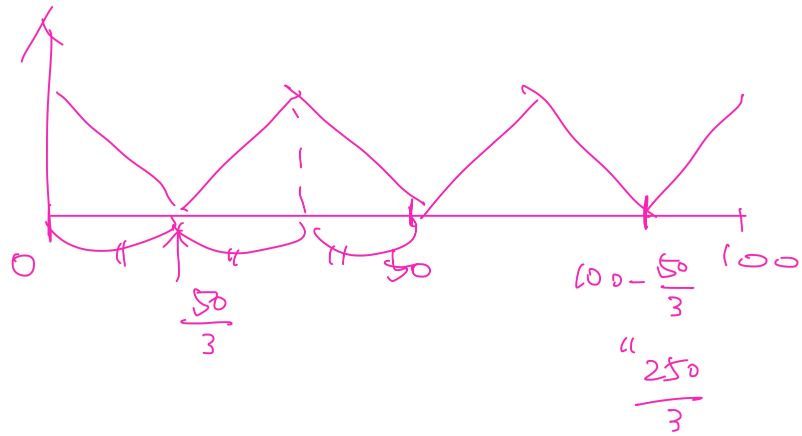
Optimal Transport.

$X \sim U(0, 100)$

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







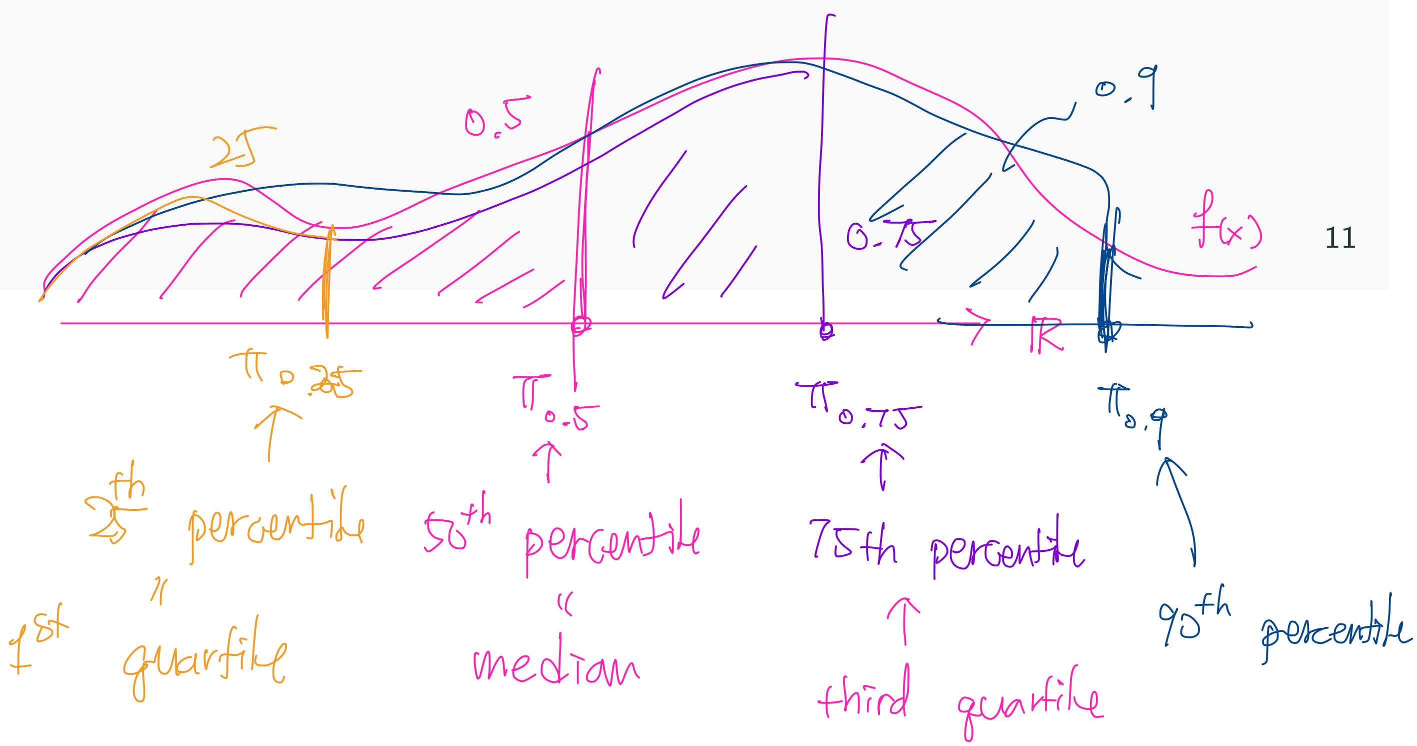
Percentile

Ex 75th percentile is $\pi_{0.75}$ s.t. $F(\pi_{0.75}) = 0.75$ ($= 100 \cdot 0.75$; 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.

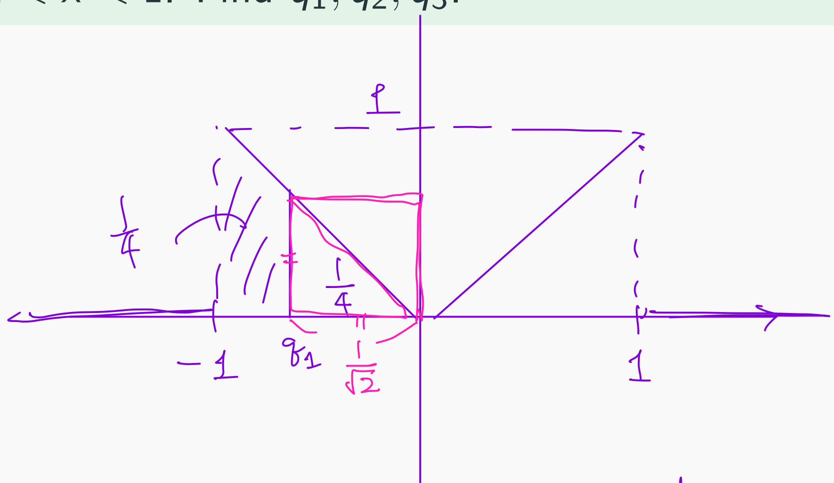
The 25th and 75th percentiles are called the first and third quartiles, respectively, and are denoted by $q_1 = \pi_{0.25}$ and $q_3 = \pi_{0.75}$.



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.

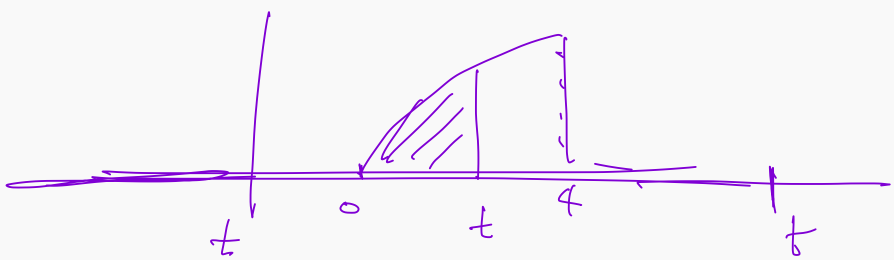


$q_2 = 50^{th}$ percentile $=$ median $= \pi_{0.5} = 0$

$q_1 = 25^{th}$ 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 $0 \le x \le 4$ be the pdf of a random variable X.

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

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

$\sqrt{x} = x^{\frac{1}{2}}$

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

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

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

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

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

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

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,

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

for $t > 0$.

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.

Exponential random variables

Theorem

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

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

$\mathrm{Var}[X] = \frac{1}{\lambda^2} = \theta^2$

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

Exponential random variables

Example

Let X have an exponential distribution with a mean $\theta = 20$.

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

Exponential random variables

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?

Gamma random variables

Consider a Poisson random variable X with λ.

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

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

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 $\alpha > 0$.

Gamma functions

The gamma function is defined by

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

for $t > 0$.

By integration by parts, we have

Gamma functions

In particular, $\Gamma(1) =$

$\Gamma(2) =$

$\Gamma(3) =$

$\Gamma(n) =$

for integers n.

Gamma random variables

Theorem

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

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

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

Gamma random variables

Example

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

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

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

Chi-square distribution

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

The pdf of X is

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

for $x > 0$.

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

Let X have an exponential distribution with mean θ.

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

Section 3.
The Normal Distribution

Gaussian random variables

Definition

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

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

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

Gaussian random variables

Theorem

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

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

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

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

Standard normal distribution

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

Example

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

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

Standard normal distribution

Theorem

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

Standard normal distribution

Example

Let $X \sim N(3, 16)$.

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

Standard normal distribution

Example

Let $X \sim N(25, 36)$.

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

Standard normal distribution

Theorem

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

Section 4.
Additional Models

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.

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

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

Weibull distribution

Definition

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

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

W is called the Weibull random variable.

Weibull distribution

Example

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

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

If W_1, W_2 are independent Weibull with α and β above, is the minimum of W_1, W_2 Weibull?

Weibull distribution

Theorem

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

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

Mixed type random variables

Example

Suppose X has a cdf

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

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

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.

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 $F(x)$.

Exercise

The cdf of X is given by

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

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