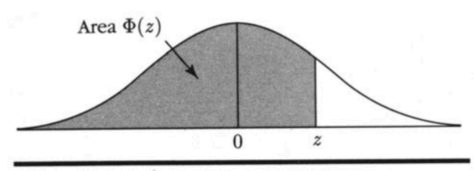
Stat 134 lec 21

Warmup 11:00–11:10

11. A large lot of marbles have diameters which are approximately normally distributed with a mean of 1 cm. One third have diameters greater than 1.1 cm. Find:

a) the standard deviation of the distribution;

You might need this:



Appendix 5
Normal Table

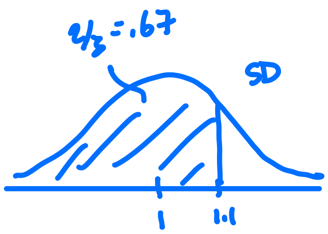


Table shows values of $\Phi(z)$ for z from 0 to 3.59 by steps of .01. Example: to find $\Phi(1.23)$, look in row 1.2 and column .03 to find $\Phi(1.2 + .03) = \Phi(1.23) = .8907$. Use $\Phi(z) = 1 - \Phi(-z)$ for negative z.

	.0	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852

From table we see

$\Phi(.44) = .67$

$.44 = \frac{1.1 - 1}{SD}$

$\Rightarrow SD = \frac{.1}{.44} = \boxed{.23}$

Last time sec 4.1 Continuous distributions

A continuous RV X, has a prob density function, $f(x)$, where $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$.

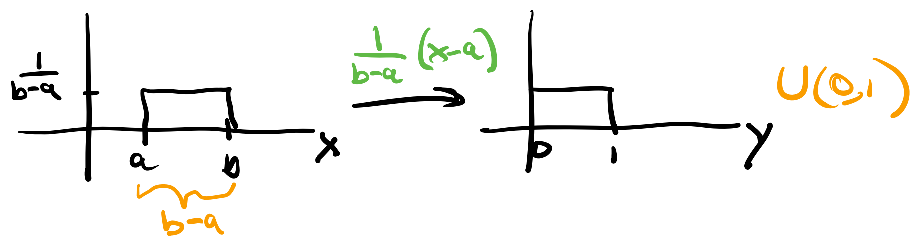
$P(X=a) = \int_a^a f(x)dx = 0$ so $P(X \geq a) = P(X > a)$.

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$$

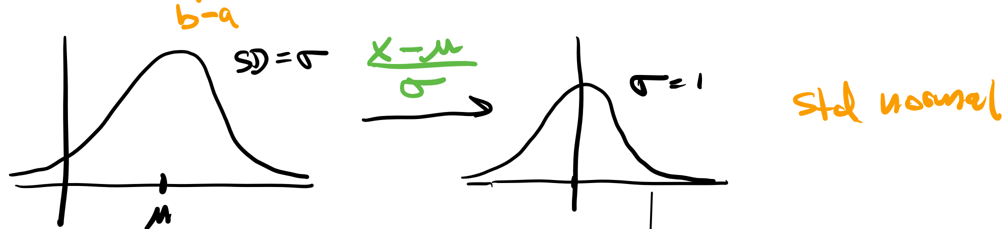
A change of scale is a transformation $Y = m + nX$, of X. The purpose is that it makes it easier to calculate $E(X)$ and $Var(X)$. It maps one density to another.

constants.

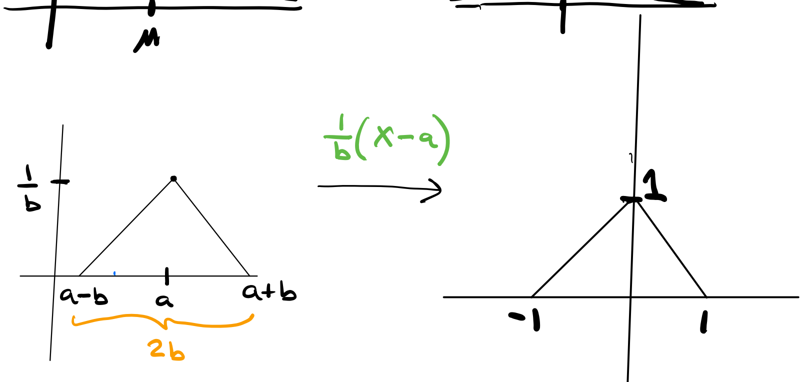
ex



ex



ex

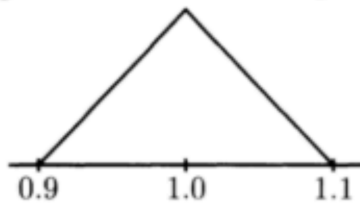


Today

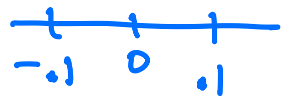
1. sec 4.1 change of scale calculations
2. briefly sec 4.5 Cummulative Distribution Function (CDF)
3. sec 4.2 Exponential Distribution.

① Sec 4.1 Change of scale calculation

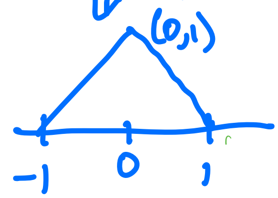
Suppose a manufacturing process designed to produce rods of length 1 inch exactly, in fact produces rods with length distributed according to the density graphed below.



subtract 1



mult by 10



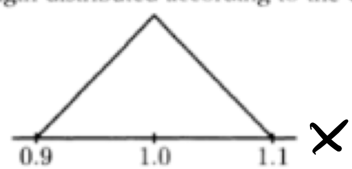
You should change the scale of X= the length of rods to:

- a: X-1
- b: .1(X-1)
- c: 10X-1
- ~~d: none of the above~~

$$\boxed{\frac{1}{1/10}(X-1)}$$

ex

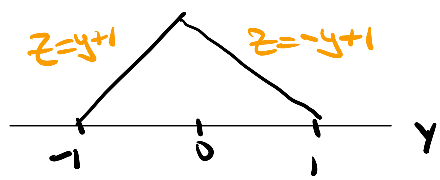
Suppose a manufacturing process designed to produce rods of length 1 inch exactly, in fact produces rods with length distributed according to the density graphed below.



Find the variance of the length of the rods.

$Y = 10(X-1)$ change of scale.

$Var(Y) = 100\,Var(X) \Rightarrow Var(X) = \dfrac{Var(Y)}{100}$ ← easier to find.



Find the density of Y:

$$f(y) = \begin{cases} y+1 & -1 \le y \le 0 \\ -y+1 & 0 \le y \le 1 \\ 0 & \text{else} \end{cases}$$

Find Var(X)

$$Var(Y) = E(Y^2) - \underbrace{(E(Y))^2}_{=0} = E(Y^2)$$

$$E(Y^2) = \int_{-\infty}^{\infty} y^2 f(y)\,dt = \int_{-1}^{0} y^2(y+1)\,dy + \int_{0}^{1} y^2(-y+1)\,dy$$

$$= \int_{-1}^{0} (y^3 + y^2)\,dy + \int_{0}^{1} (-y^3 + y^2)\,dy = \frac{y^4}{4} + \frac{y^3}{3}\Big|_{-1}^{0} + \frac{-y^4}{4} + \frac{y^3}{3}\Big|_{0}^{1} = \boxed{1/6}$$

$$\Rightarrow Var(Y) = 1/6 \Rightarrow Var(X) = \frac{Var(Y)}{100} = \boxed{1/600}$$

② briefly sec 4.5 The Cummulative Distribution Function (CDF)

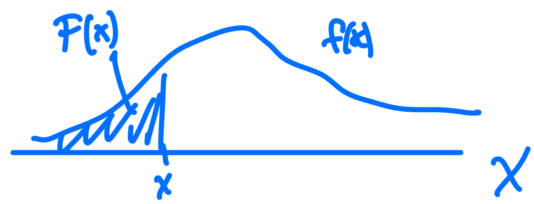
Let X be a continuous RV

$F(x) = P(X \le x)$ → a number between 0 and 1

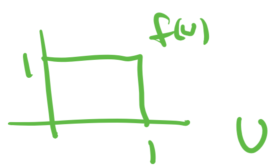
If $f(x)$ is a density of X,

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)\,dx$$

Picture

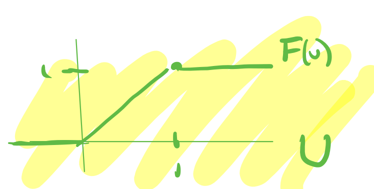


ex $U \sim \text{Unif}(0,1)$



$$f(u) = \begin{cases} 1 & 0 < u < 1 \\ 0 & \text{else} \end{cases}$$

$$F(u) = \int_0^u 1\,dx = u$$



$$\boxed{F(u) = \begin{cases} 0 & -\infty < u \le 0 \\ u & 0 \le u \le 1 \\ 1 & u \ge 1 \end{cases}}$$

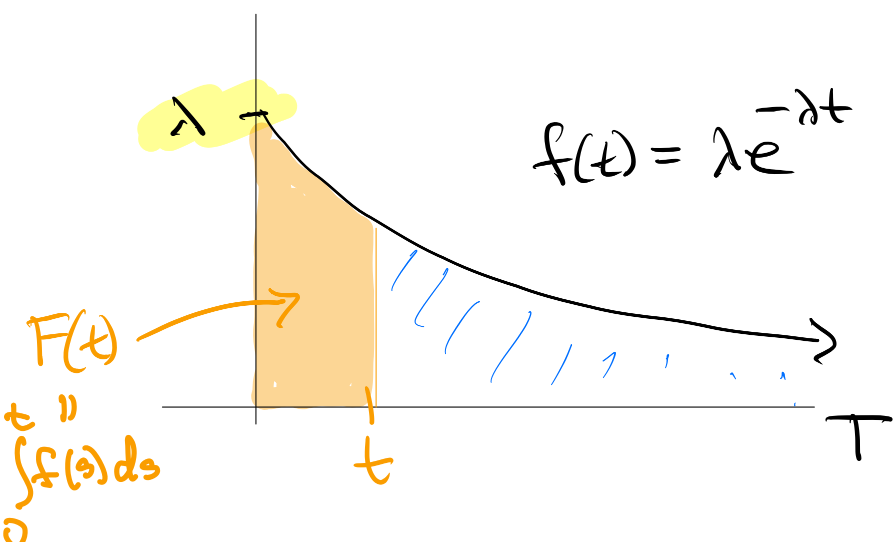
By FTC, $F'(x) = f(x)$

Consequently a density function and cdf are equivalent descriptions of a RV.

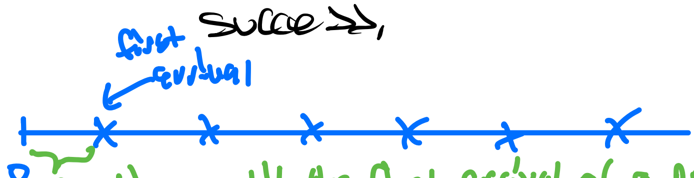
③ Sec 4.2   Exponential distribution

Def^n A random time T has exponential distribution with rate $\lambda > 0$.

$T \sim Exp(\lambda)$, if T has density $f(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & \text{else} \end{cases}$



ex T = time until your first success where λ = rate of success,



time until the first arrival of a poisson process with rate λ.

ex T = time until a lightbulb burns out

CDF and survival function

$T \sim Exp(\lambda) \quad f(t) = \lambda e^{-\lambda t}$

Compute the CDF of T.

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

$$= \int_{-\infty}^{0} \underbrace{f(s)}_{=0}\,ds + \int_{0}^{t} f(s)\,ds = \int_{0}^{t} \lambda e^{-\lambda s}\,ds = \frac{\cancel{\lambda} e^{-\lambda s}}{-\cancel{\lambda}}\Big|_{0}^{t}$$

$$= -e^{-\lambda t} + 1 = \boxed{1 - e^{-\lambda t}}$$

$P(T > t) = e^{-\lambda t}$ is called the <u>survival function</u>

$T \sim Exp(\lambda)$ iff $P(T > t) = e^{-\lambda t}$

since $\overset{= 1 - P(T > t)}{F(t)}$ and $f(t)$ both define distribution.

ex

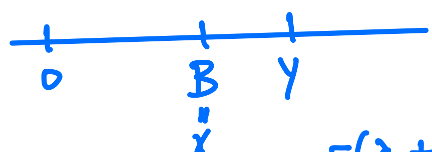
GSI Brian and Yiming are each helping a student. Brian and Yiming see students at a rate λ_B and λ_Y students per hour respectively.

let B = wait time for Brian $\sim \text{Exp}(\lambda_B)$

Y = wait time for Yiming $\sim \text{Exp}(\lambda_Y)$

What distribution is $X = \min(B, Y)$?

$P(X > x)$



$=$

$P(B > x, Y > x) = P(B > x)\,P(Y > x) = e^{-(\lambda_B + \lambda_Y)x}$

$P(B > x) = e^{-\lambda_B x}$, $\quad P(Y > x) = e^{-\lambda_Y x}$

$\Rightarrow X \sim \text{Exp}(\lambda_B + \lambda_Y)$

Memoryless property of Exponential

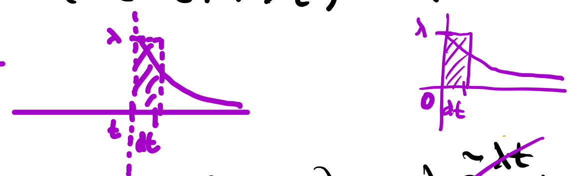
survival function

$T \sim \text{Exp}(\lambda)$ $\quad f(t) = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & \text{else} \end{cases}$ $\quad$ or $\quad P(T>t) = e^{-\lambda t}$

T = time until 1st success (arrival)

Memoryless Property : $P(T \in dt \mid T > t) = P(0 < T < 0 + dt)$

picture



$$P(T \in dt \mid T > t) = \frac{P(T \in dt)}{P(T > t)} \approx \frac{\lambda \cancel{e^{-\lambda t}}\, dt}{\cancel{e^{-\lambda t}}} = \boxed{\lambda dt}$$

$\parallel$

$P(0 < T < 0 + dt)$

another picture

conclusions:

① So long as a lightbulb is still functioning it is as good as new

② $P(T \in dt \mid T > t) = \lambda dt \Rightarrow \quad \lambda = \dfrac{P(T \in dt \mid T > t)}{dt}$

the instantaneous rate of success is the same for all t

Memoryless property of the geometric distribution

Let $X \sim \text{Geom}(p)$

$P(X > 3) = ?$ q^3

Question

Given it takes you more than $j = 10$ p-coin tosses to get your first heads, what is the chance it will take you more than $k + j = 13$ coin tosses to get your first heads? — q^3

Says, the geom distribution has no memory what happened before the present time

proved in appendix of notes

Thm The Geometric distribution is the only discrete distribution with values 1, 2, 3, ... having the memoryless property

$$P(X > k + j \mid X > j) = P(X > k)$$

Only 2 distributions are memoryless:

For discrete $(X=1,2,3,\dots)$ — Geometric

For continuous $(T>0)$ — Exponential

Proof is similar to geom.

Appendix

First we show that $P(X > k) = q^k$ is an equivalent def^n of Geometric on $1, 2, 3, \dots$

let $X \sim \text{Geom}(p)$ where

$X = \#$ trials until your first success.

$$P(X=k) = q^{k-1}p$$

recall
$$P(X>k) = P(X=k+1) + P(X=k+2) + \dots$$
$$= q^k p + q^{k+1}p + \dots$$
$$= q^k p \underbrace{(1 + q + q^2 + \dots)}_{= \frac{1}{1-q}}$$
$$= q^k.$$

Since
$$P(X=k) = \underbrace{P(X>k-1)}_{= q^{k-1}} - \underbrace{P(X>k)}_{= q^k}$$
$$= q^{k-1}(1-q) = q^{k-1}p$$

$P(X>k) = q^k \Rightarrow X$ is Geom(p).

$\Rightarrow$ $X \sim \text{Geom}(p)$ iff $P(X>k) = q^k$

<u>Thm</u> The Geometric distribution is the only discrete distribution with values 1, 2, 3, ... having the

<u>memoryless property</u>

$$P(X > k+j \mid X > j) = P(X > k)$$

Proof/

Let $X \sim \text{Geom}(p)$, $X = 1, 2, 3, \ldots$

$$P(X > k+j \mid X > j) = \frac{P(X > k+j,\ X > j)}{P(X > j)}$$

$$= \frac{P(X > k+j)}{P(X > j)}$$

$$= \frac{q^{k+j}}{q^{j}} = q^{k}$$

$= P(X > k)$ ✓

Conversely,

For positive integers k, j

Suppose $P(X > k+j \mid X > j) = P(X > k)$

$$\begin{gathered} P(X > k+j \mid X > j) \\ \parallel \\ \frac{P(X > k+j)}{P(X > j)} \end{gathered}$$

$\iff P(X > k+j) = P(X > j)P(X > k)$

let $q = P(X > 1)$

Show that $P(X > j) = q^j$

for $j = 1, 2, 3, \ldots$ by induction.

Base case:

$$P(X>1) = q^1 \quad \checkmark$$

Assume $P(X>j-1) = q^{j-1}$

$$P(X>j) = P(X>j-1+1)$$

$$= \underbrace{P(X>j-1)}_{=\,q^{j-1}}\underbrace{P(X>1)}_{=\,q} = q^j$$

$\Rightarrow X \sim \text{Geom}(p)$. $\square$