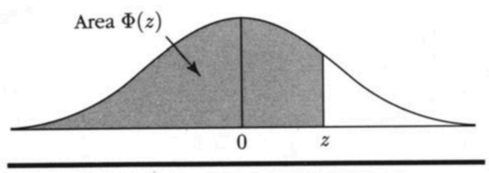
Stat 134 lec 20

Warmup

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

Change of scale

$$z = \frac{X - E(X)}{\sigma}$$

$\Phi(.44) = .67$ from table

so $z = .44$

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

$$\to SD(X) = \frac{.1}{.44} = \boxed{.23}$$

Last time sec 4.1 Continuous distributions

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

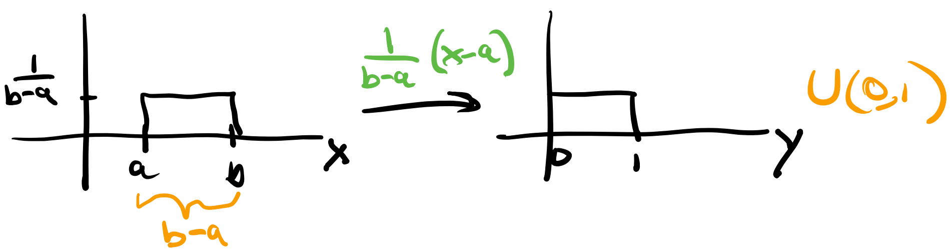
$P(X=a) = \int_a^a f(x)dx = 0$ so $P(X \geq a) = P(X > a)$.

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

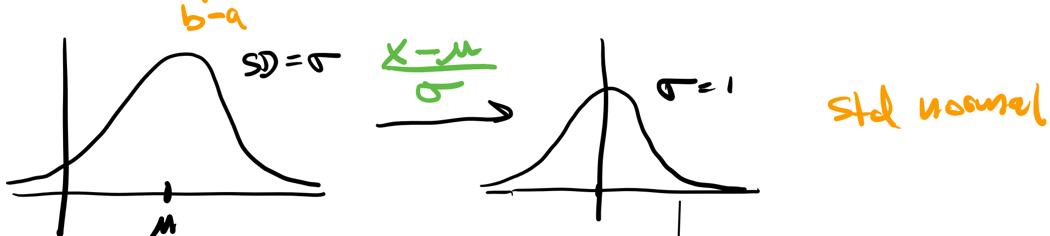
A change of scale is a transformation $Y = m + nX$, of X. The purpose is that it makes it easier to calculate $E(X)$ and $Var(X)$. It maps one density to another.

constants.

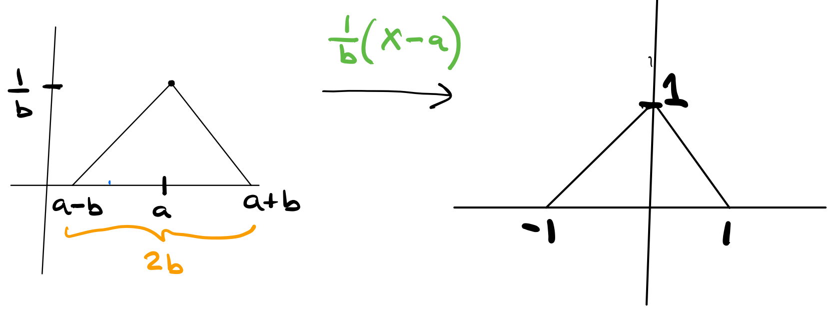
ex



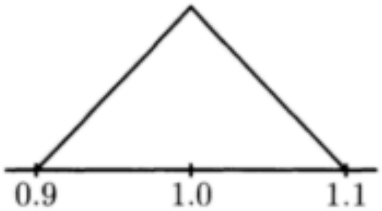
ex



ex



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.



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

a: X-1	The area is already 1

d: none of the above	You need to move the center of the rod from x = 1 to x = 0 so first we need to subtract 1. Now we have the end point at x = .1 and to normalize this so the end point is at x = 1 we need to divide by .1 which is equivalent to multiplying by 10. So our final equation is 10(x - 1)

<u>Today</u>

① briefly sec 4.5 Cummulative Distribution Function (CDF)

② Sec 4.2 Exponential Distribution.

① briefly sec 4.5 The Cummulative Distribution Function (CDF)

Let X be a continuous RV

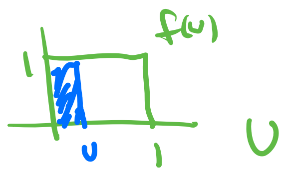
$F(x) = P(X \leq x)$ — a number between 0 and 1

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

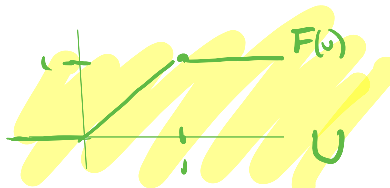
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx$$

ex $U \sim Unif(0,1)$

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



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



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

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

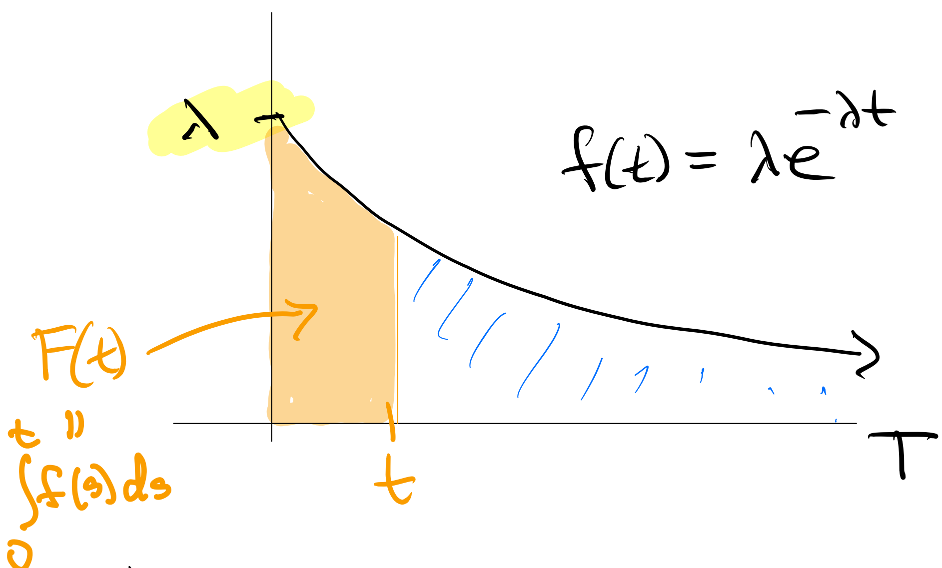
Consequently a density function and cdf are equivalent descriptions of a RV.

② Sec 4.2 — Exponential distribution

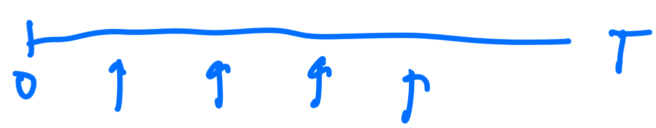
Def^n A random time T has exponential distribution with rate $\lambda > 0$.

$T \sim Exp(\lambda)$, 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.



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.

$$F(t) = P(T \leq t) = \int_0^t \lambda e^{-\lambda s}\, ds = \frac{\lambda e^{-\lambda s}}{-\lambda}\Big|_0^t$$

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

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

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

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

tinyurl.com/mav6-2023

Stat 134

1. GSI Brian and Yiming are each helping a student. Brian and Yiming see students at a rate of λ_B and λ_Y students per hour respectively.
   Let
   $$B = \text{wait time for Brian } \sim Exp(\lambda_B)$$
   $$Y = \text{wait time for Yiming } \sim Exp(\lambda_Y)$$
   What distribution is $T = \min(B, Y)$?
   Hint: compute $P(T > t)$

   a $Exp(\max(\lambda_B, \lambda_Y))$

   b $Exp(\lambda_B - \lambda_Y)$

   c $Exp(\lambda_B + \lambda_Y)$

   d none of the above

$$P(T>t) = \underset{\overset{\parallel}{e^{-\lambda_B t}}}{P(B>t)}\,\underset{\overset{\parallel}{e^{-\lambda_y t}}}{P(y>t)} = e^{-(\lambda_B+\lambda_y)t}$$

$$\Rightarrow \boxed{T \sim Exp(\lambda_B+\lambda_y)}$$

The memoryless property

This property relates to the geometric and exponential distributions.

In words, it says if you haven't had success yet then you can reset the clock to zero.

More formally,

if $T \sim \text{Exp}(\lambda)$, $T=$ time until your first success or arrival.

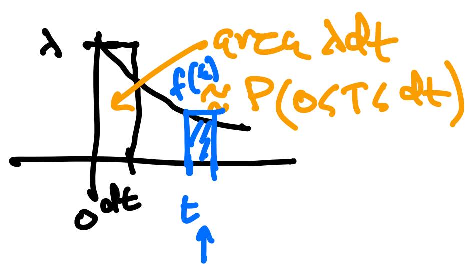
the memoryless property says:

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

you have success in small interval after time t

given you haven't had success before time t

you have success in small time interval after time 0.



By the graph of $T \sim \text{Exp}(\lambda)$, we see $P(0 < T < dt) \approx \lambda dt$

By Baye's rule $P(T \in dt \mid T > t) = \dfrac{P(T \in dt,\ T > t)}{P(T > t)}$

$$= \frac{P(T \in dt)}{P(T > t)}$$

$$\approx \frac{f(t)\,dt}{e^{-\lambda t}} = \frac{\lambda \cancel{e^{-\lambda t}}\,dt}{\cancel{e^{-\lambda t}}}$$

$$= \lambda\,dt$$

$$\approx P(0 \leq T < dt)$$

Proving the memoryless property of exponential.

Interestingly,

Only 2 distributions are memoryless:

For discrete $(X = 1, 2, 3, \ldots)$ – Geometric (see lec 14)

For continuous $(T > 0)$ – Exponential

ex

A family is getting ready for their trip to Yosemite. Each person is in their room, packing their bags. For each person, the time it takes them to pack their bag is exponentially distributed and independent of the time it takes any other person. On average, it takes each parent 1 hour and each child 2 hours to get ready. In a family with 2 parents and 4 children, what is the probability that it takes the family more than 2 hours to get ready?

let $P_1, P_2, C_3, C_4, C_5, C_6$ be the packing time of the parents and children.

If a child takes 2 hrs on average to pack what is his average packing rate in units of $\frac{1}{hr}$? — answ $\frac{1}{2} hr^{-1}$

(in other words they are half packed in an hour.)

The average packing rate is the same as the rate since we assume λ never changes in a Poisson process.

Hence $P_1, P_2 \overset{iid}{\sim} \text{Exp}(1)$

$C_1, C_2, C_3, C_4 \overset{iid}{\sim} \text{Exp}(1/2)$

$M = \max(P_1, P_2, C_3, C_4, C_5, C_6)$

$P(M > 2)$ is the probability at least one person isn't packed in 2 hrs.

$P(M > 2) = 1 - P(M \le 2)$ ← probability everyone is packed in 2 hrs.

$$= 1 - \underbrace{P(P_1 \le 2)}_{1-e^{-1\cdot 2}}\underbrace{P(P_2 \le 2)}_{1-e^{-1\cdot 2}}\underbrace{P(C_1 \le 2)}_{1-e^{-\frac{1}{2}\cdot 2}} \cdots \underbrace{P(C_6 \le 2)}_{1-e^{-\frac{1}{2}\cdot 2}}$$

$$= \boxed{1 - (1-e^{-2})^2(1-e^{-1})^4}$$