Stat 134 lec 21

Last time sec 4.1 Continuous distributions

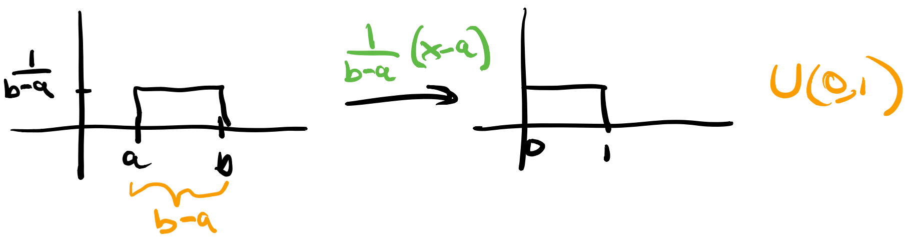
A continuous RV X, has a prob density function, $f(x)$, where $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$.

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

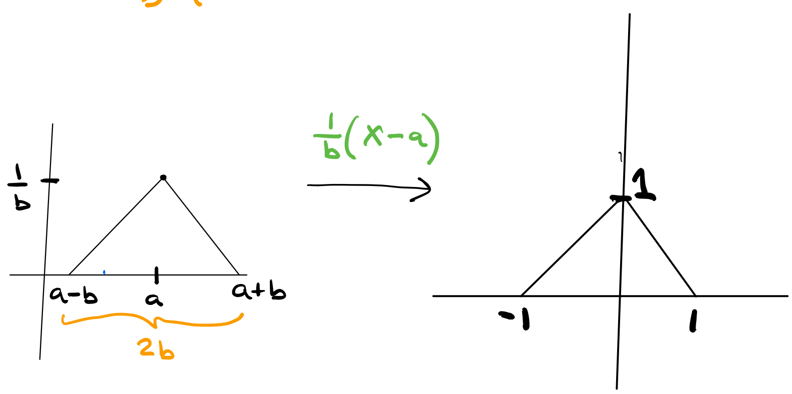
A change of scale is a transformation $Y = c + dX$, of X. The density of X gets transformed into the density of Y.

constants.

ex



ex

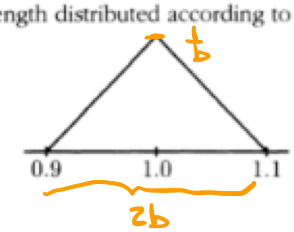


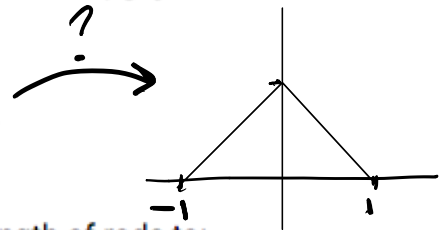
Today
1. review concept test responses from last time.
2. sec 4.1 change of scale calculations
3. briefly sec 4.5 Cummulative Distribution Function (CDF)
4. Sec 4.2 Exponential Distribution.

(1) student responses.

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.

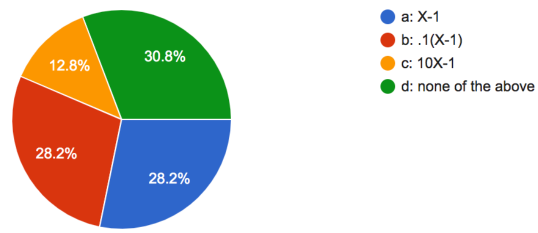
a = 1
b = 1/10





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



b: .1(X-1)	To make the triangle into standard normal, translate it so that 1.0 becomes 0, and 0.9 becomes -0.1 and 1.1 becomes 0.1. This is the translation of X-1. To make the area equal to 1, divide height by 10 which is b.

1/10

10

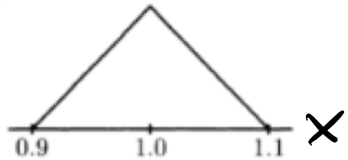
a: X-1	Height looks to be 1, so we do not need to change that. Only need to center vertex at 0 by subtracting 1.

d: none of the above	The graph above is a density so the area under the curve has to be 1. Transfer the tip to the origin, so it's X - 1. The height should be 1 so multiply 10 by X - 1

② Sec 4.1 Change of scale calculation

ex

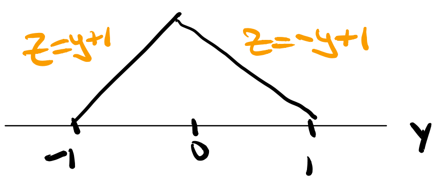
Suppose a manufacturing process designed to produce rods of length 1 inch exactly, in fact produces rods with length distributed according to the density graphed below.



Find the variance of the length of the rods.

$Y = 10(X-1)$ change of scale. ← easier to find.

$Var(Y) = 100\, Var(X) \Rightarrow Var(X) = \dfrac{Var(Y)}{100}$



Find Var(X)

$Var(Y) = E(Y^2) - E(Y)^2$

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

$$= \int_{-1}^{0} (y^3+y^2)\,dy + \int_{0}^{1} (-y^3+y^2)\,dy = \left.\frac{y^4}{4} + \frac{y^3}{3}\right|_{-1}^{0} + \left.-\frac{y^4}{4} + \frac{y^3}{3}\right|_{0}^{1}$$

$$= -\frac{1}{4} + \frac{1}{3} + -\frac{1}{4} + \frac{1}{3} = -\frac{1}{2} + \frac{2}{3} = \boxed{\frac{1}{6}}$$

$E(Y) = 0$

$\Rightarrow Var(Y) = \boxed{\frac{1}{6}}$

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

③ briefly sec 4.5 The Cummulative Distribution Function (CDF)

Let X be a continuous RV

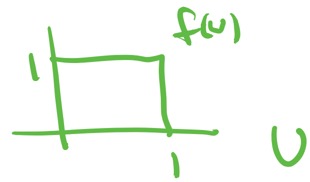
$F(x) = P(X \leq x)$ → a number between 0 and 1

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

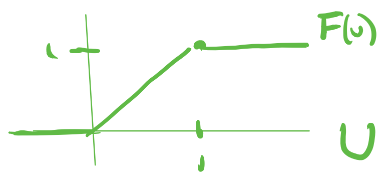
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx$$

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

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



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



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

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

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

④ Sec 4.2 Exponential Distribution

Memoryless property of the geometric distribution

$X \sim Geom(p)$

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

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

recall $P(X>k) = P(X=k+1) + P(X=k+2) + \cdots$

$= q^k p + q^{k+1} p + \cdots$

$= q^k p \left(1 + q + q^2 + \cdots\right)$

$= q^k \cdot \qquad \frac{1}{1-q}$

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 $Geom(p)$.

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

Question

If 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?

answer

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

Think of starting at $j+1^{st} = 11^{th}$ toss

so it should be $P(x>3) = q^3$

i.e $P(\underbrace{X > k+j}_{=13} \mid \underbrace{X > j}_{=10}) = P(X > \underbrace{k}_{=3})$

Thm The Geometric distribution is the only discrete distribution with values $1, 2, 3, \ldots$ having the

memoryless property

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

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

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

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

let $q = P(X > 1)$

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

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

Base case:

$P(X>1)=q^1$ ✓

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

Finish the proof.

$P(X>j)=P(X>\underbrace{j-1}+1)$

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

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

Only 2 distributions are memoryless:

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

For continuous $(T>0)$ – Exponential

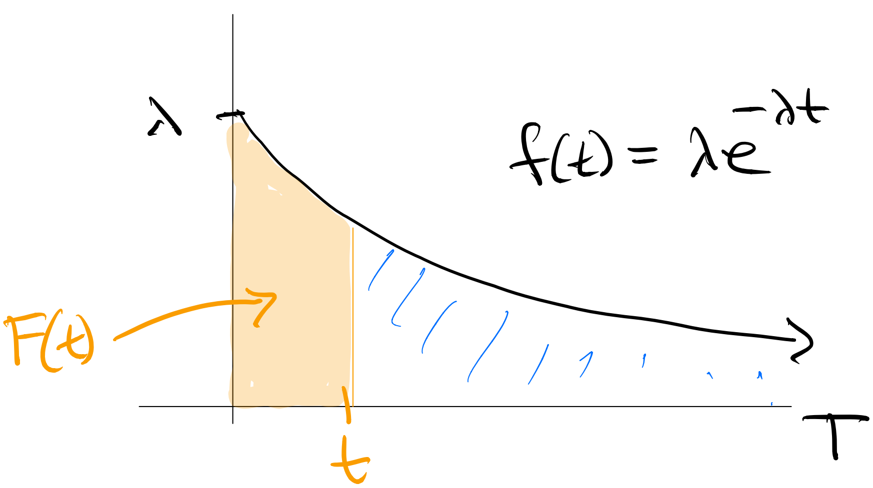
Proof is similar to geom.

Exponential distribution

<u>Def</u>ⁿ

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

$T \sim Exp(\lambda)$, if T has density $f(t)=\lambda e^{-\lambda t}$ $t \geq 0$



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

ex T = time until a lightbulb burns out

CDF and survival function

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

Compute the CDF of T.

$$F(t) = P(T \le t) = \int_0^t f(s)\,ds$$

$$= \int_0^t \lambda e^{-\lambda s}\,ds = \left.\frac{\lambda e^{-\lambda s}}{-\lambda}\right|_0^t$$

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

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

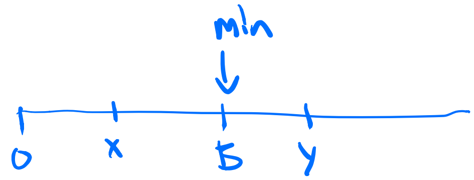
$T \sim \text{Exp}(\lambda)$ iff $P(T > t) = e^{-\lambda t}$

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

ex Brian and Yiming arrive at the beginning of a group advising session. The amount of time they will stay is exponentially distributed with rates λ_B and λ_Y, independent of each other.

$$\left(\text{i.e. } \left.\begin{aligned} B &= \text{Brian's wait} \sim \text{Exp}(\lambda_B) \\ Y &= \text{Yiming's wait} \sim \text{Exp}(\lambda_Y) \end{aligned}\right\} \text{indep.}\right)$$

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



$\leftarrow$ If $X > x$ then $B > x$ and $Y > x$

$\leftarrow \min(B,Y)$

$$P(X > x) = P(B > x, Y > x) = \underset{\substack{\| \\ e^{-\lambda_B x}}}{P(X > x)}\ \underset{\substack{\| \\ e^{-\lambda_Y x}}}{P(Y > y)}$$

$$= e^{-(\lambda_B + \lambda_Y)x}$$

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