Stat 134 lec 20 (no lec 19)

last time midterm! Nice Job!!

today Sec 4.1 Continuous distributions
- prob density
- expectation and variance

Sec 4.1 Probability density.

let X be a continuous RV.

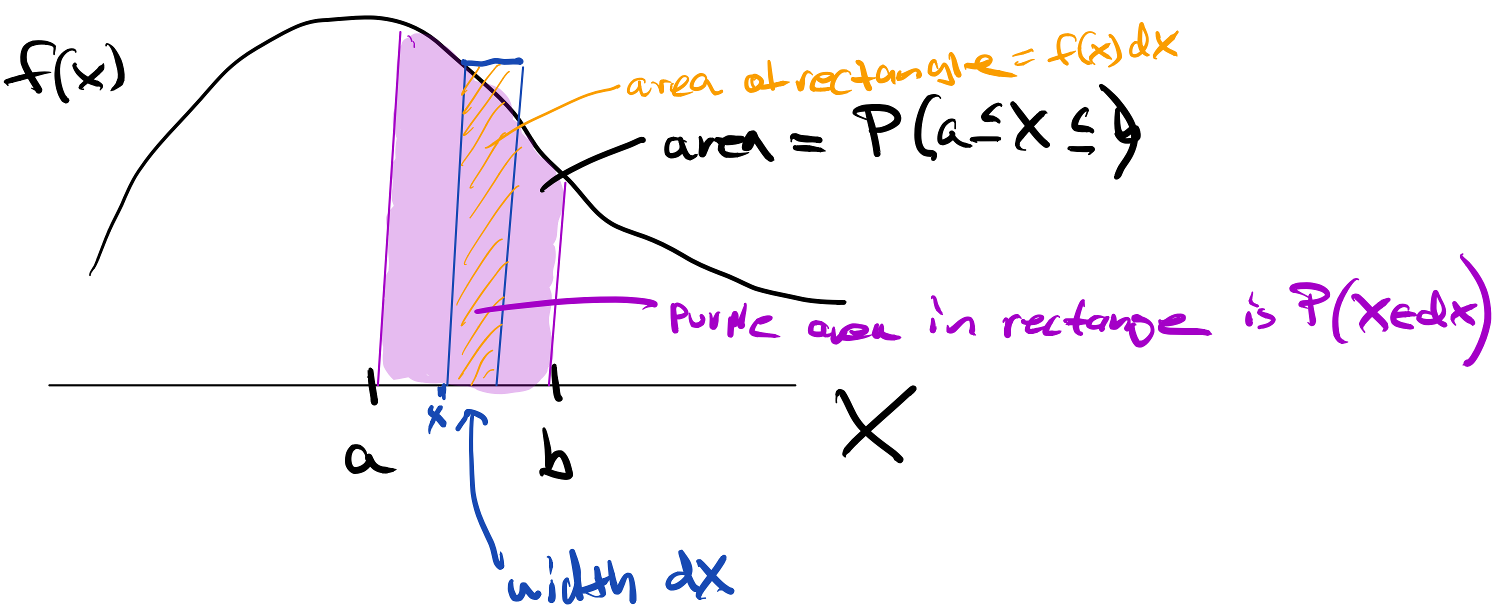
The prob density of X is described by a **prob density function** $f(x) \geq 0$ for $x \in X$

and $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$

ex For the std. normal distribution,

$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$ is a prob density function

Picture:



we see from the rectangle in the picture,

$P(X \in dx) \approx f(x)dx$ (notice purple and orange area not same)

here dx = tiny interval around x and also the length of the interval

For all a, b $P(a \le X \le b) = \int_a^b P(X \in dx)dx = \int_a^b f(x)dx$

Note $f(x)$ is not a probability.

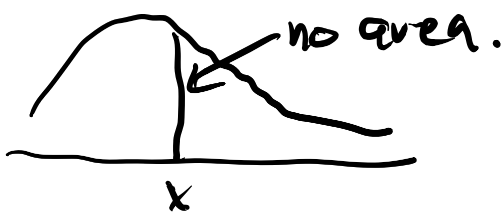
$f(x)dx$ is a probability

$$f(x) \approx \frac{P(X \in dx)}{dx}$$

units of f? — Prob per unit length

hence "Prob density"

$P(X=x) = 0$

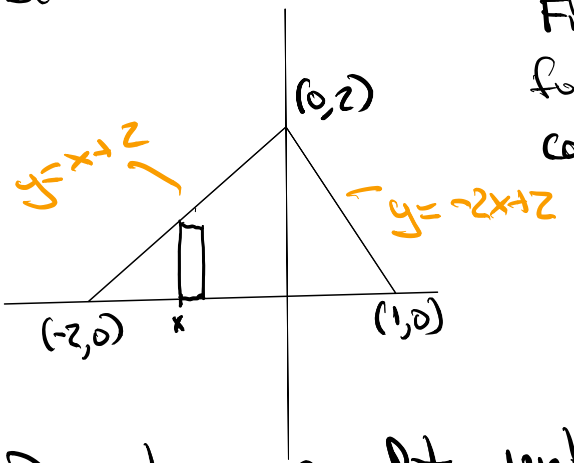


Hence $P(a \le X \le b) = P(a < X < b)$

(we don't have to worry about endpts).

ex 4.1.12b

Consider a point picked uniformly at random from the area inside the shape below.



Find the density function of the x coordinate.

Drawing a pt uniformly at random means that you are throwing a dart at the shape. All pts in the rectangle drawn above have x coordinate x. The probability of getting an x coordinate of x is $P(X \in dx)$.

$P(x \in dx)$ must integrate to 1. For this to happen we must have:

$$P(X \in dx) \approx \begin{cases} \frac{(x+2)dx}{3} & \text{if } -2 \le x \le 0 \\ \frac{(-2x+2)dx}{3} & \text{if } 0 \le x \le 1 \\ 0 & \text{else} \end{cases}$$

Since the total area of the triangle is 3.

Since $P(X \in dx) \approx f(x)dx$ it follows that,

$$f(x) = \begin{cases} \frac{x+2}{3} & \text{if } -2 \le x \le 0 \\ \frac{-2x+2}{3} & \text{if } 0 \le x \le 1 \\ 0 & \text{else} \end{cases}$$

is a density function.

Expectation and Variance

For discrete,

$$E(g(X)) = \sum_{x \in X} g(x) P(X=x)$$

(assuming $E(|g(X)|) < \infty$)

For continuous,

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

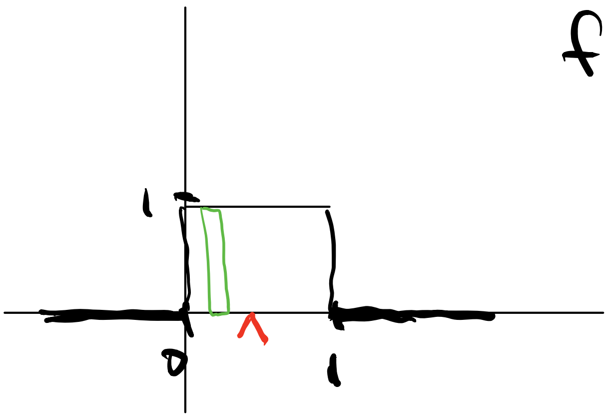
(assuming $E(|g(X)|) < \infty$)

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

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

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

ex U : unif$(0,1)$ — standard uniform



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

$$E(U)=\int_0^1 x\,dx=\frac{x^2}{2}\Big|_0^1=\frac{1}{2}$$

$$E(U^2)=\int_0^1 x^2\,dx=\frac{x^3}{3}\Big|_0^1=\frac{1}{3}$$

$$\text{Var}(U)=\frac{1}{3}-\left(\frac{1}{2}\right)^2=\frac{1}{12}$$

Sometimes to calculate $E(X)$, $\text{Var}(X)$ and probabilities it is convenient to make a <u>change of scale</u> (i.e convert X to $cX+d$ for constants c, d.

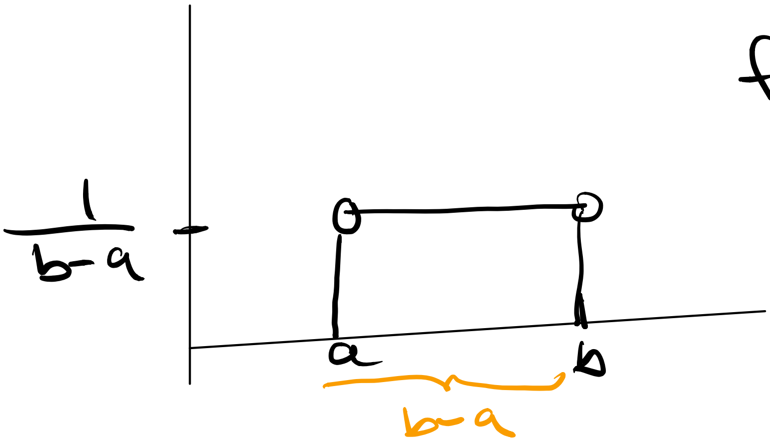
<u>Change of scale</u> — $cX+d$ for constants c, d.

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

We define $Z=\dfrac{X-\mu}{\sigma}=\underbrace{\frac{1}{\sigma}}_{c}X+\underbrace{\left(-\frac{\mu}{\sigma}\right)}_{d}$.

Z is the standard normal RV and is easier to work with.

ex: let $X \sim \text{Unif}(a,b)$



$$f(x) = \begin{cases} \frac{1}{b-a}, & a<x<b \\ 0, & \text{else} \end{cases}$$

Find $E(X)$, $Var(X)$.

idea: Relate X to U by a change of scale.

$$U = \frac{X-a}{b-a} = \underbrace{\frac{1}{b-a}}_{= c} X + \underbrace{\left(-\frac{a}{b-a}\right)}_{= d}$$

is a change of scale of X.

$$\Rightarrow \quad X = (b-a)U + a$$

Then $E(X) = (b-a)\underbrace{E(U)}_{= 1/2} + a$

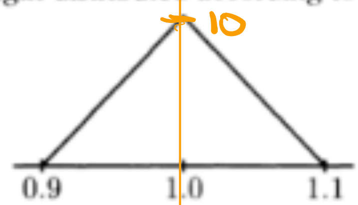
$$= \boxed{\frac{b+a}{2}}$$

$$Var(X) = Var\left((b-a)U + a\right)$$

$$= (b-a)^2 \underbrace{Var(U)}_{= 1/12} = \boxed{\frac{(b-a)^2}{12}}$$

Concept test

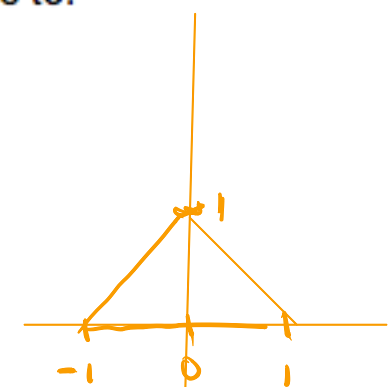
Suppose a manufacturing process designed to produce rods of length 1 inch exactly, in fact produces rods with length distributed according to the density graphed below.



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

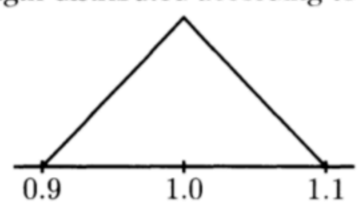
$Y = 10(X-1) \rightsquigarrow$



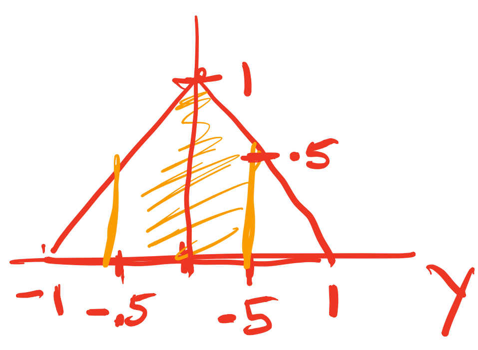
$$f_Y(y) = \begin{cases} -y+1 & \text{if } 0 \le y \le 1 \\ y+1 & \text{if } -1 \le y \le 0 \\ 0 & \text{else} \end{cases}$$

Concept test pt 2.

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.



Explain in detail how to find the proportion of rods between .95 and 1.05?



<u>Soln</u>

$$Y = 10(X-1)$$

$$X = \frac{Y}{10} + 1$$

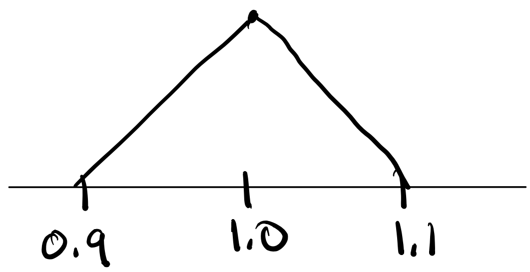
$$P(.95 < X < 1.05) = P\left(.95 < \frac{Y}{10} + 1 < 1.05\right)$$

$$= P(-.5 < Y < .5)$$

$$= 1 - (.5)^2 = .75$$

ex

Suppose a manufacturing process designed to produce rods of length 1 inch exactly has distribution with density:



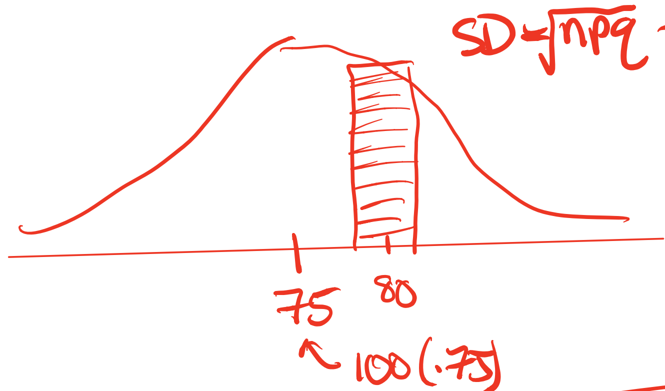
a) What is the chance that exactly 80 out of 100 rods are between .95 and 1.05?

soln

$P = .75$

$$\binom{100}{80}(.75)^{80}(.25)^{20}$$

$SD = \sqrt{npq} = \sqrt{100(.75)(.25)}$
$= \sqrt{18.75}$



$$\approx \boxed{\Phi\left(\frac{80.5-75}{\sqrt{18.75}}\right) - \Phi\left(\frac{79.5-75}{\sqrt{18.75}}\right)}$$

b) Find the variance of the length of the rods.

soln

$$Y = 10(X-1)$$

$$\text{var}(Y) = 100\,\text{var}(X).$$

$$\Rightarrow \text{var}(X) = \frac{\text{var}(Y)}{100}.$$

$$\text{var}(Y) = E(Y^2) - \underset{0}{\underset{\parallel}{E(Y)}}$$

$$E(Y^2) = 2\int_0^1 y^2(1-y)\,dy = 2\left(\frac{y^3}{3} - \frac{y^4}{4}\right)\Big|_0^1$$

density even function

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

$$\Rightarrow \text{var}(Y) = \frac{1}{6} \Rightarrow \text{var}(X) = \boxed{\frac{1}{600}}$$