Stat 134 lec 20 (no lec 19)

See website for updated quiz dates
Quiz 4 Monday October 28

last time — Great job on the midterm!

comments from students last semester:

> The later parts of the course covering new, difficult material was taken at seemingly the same pace as the earlier parts covering reviewed concepts or familiar material. As time went on it became harder and harder to keep up with the course – I didn't feel I had enough time to learn one idea before we were on to the next.

> Expect to increase the amount of time you spend on this course as the semester goes on; the increase in difficulty over time is steeper than for some other courses.

> Expect to increase the amount of time you spend on this course as the semester goes on; the increase in difficulty over time is steeper than for some other courses.

> read the book ahead of time

today

Sec 4.1 Continuous distributions

1. probability density
2. expectation and variance.
3. Change of scale

① <u>sec 4.1</u> Probability density.

let X be a continuous RV

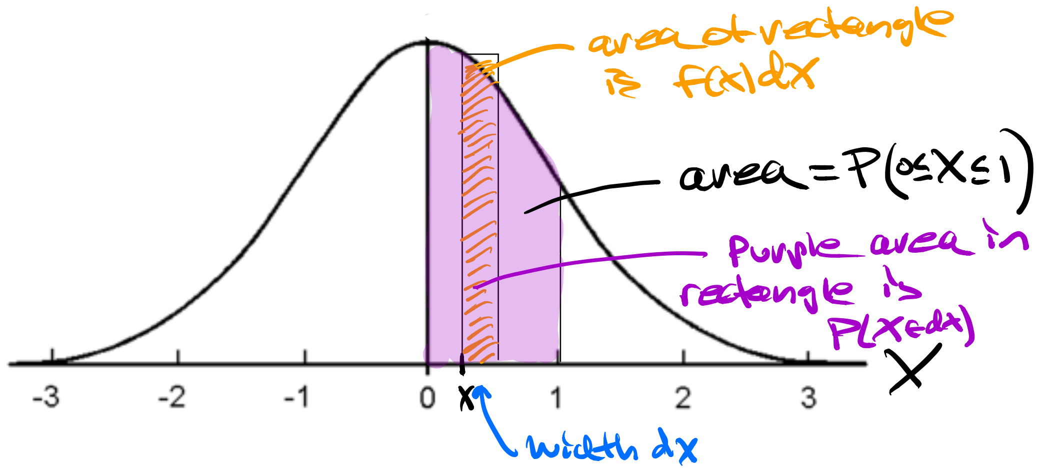
The probability density of X is described by a <u>prob density function</u>

$$f(x) \geq 0 \quad \text{for } x \in X$$

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

<u>ex</u> The <u>standard normal distribution</u>

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



Throwing a dart randomly at the histogram the x coordinate of your dart is most likely to be near zero.

The probability of getting an x coordinate of x is written $P(X \in dx)$.

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

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

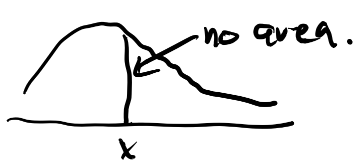
$f(x)dx$ is a probability

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

units of f? — Prob per unit length

hence "Prob density"

$P(X = x) = 0$

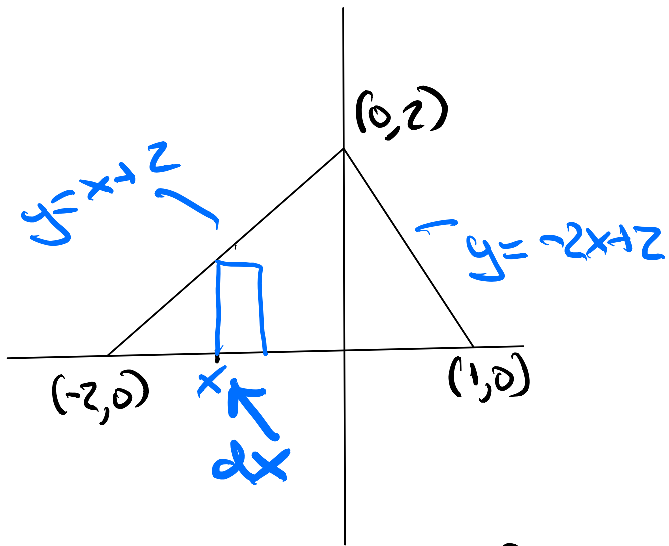


Hence $P(a \le X \le b) = P(a < X < b)$

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

ex 4.1.12 b

Consider a point picked uniformly at random from the area inside the following triangle



Find the density function of the x-coordinate $f(x)$

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

Procedure:

(1) compute area A of shape

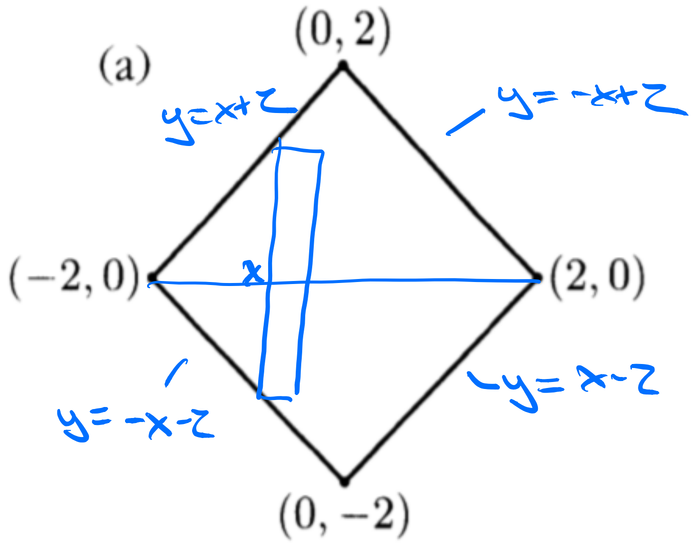
(2) for a strip width dx find $P(x \in dx) = \frac{g(x)}{A}$

Area $= \frac{1}{2}$ height $\cdot$ width
$= \frac{1}{2} 2 \cdot 3 = 3 \leftarrow$ not 1.

change vertex (0,2) to $(0, \frac{2}{3})$

ex 4.1.12 a

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



$$A = 2\cdot\left(\tfrac{1}{2}\,4\cdot 2\right) = 8$$

$$\frac{(x+2)}{4}$$

$$P(X \in dx) = \frac{2(x+2)\,dx}{8} \quad \text{for } -2 \le x \le 0$$

$$P(X \in dx) = \frac{2(-x+2)\,dx}{8} \quad \text{for } 0 \le x \le 2$$

$$\boxed{f(x) = \begin{cases} \dfrac{x+2}{4} & -2 \le x \le 0 \\[6pt] \dfrac{-x+2}{4} & 0 \le x \le 2 \end{cases}}$$

(2)

Expectation and Variance

For discrete,

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

absolutely convergent sums have nice properties we need (see lec 15)

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

For continuous,

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

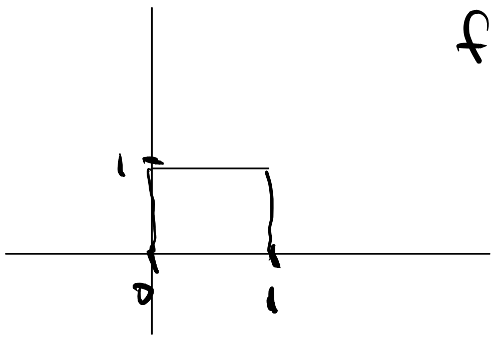
(assuming $E(|g(X)|) < \infty$)

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

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

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

ex $U : \text{unif}(0,1)$ — standard uniform



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

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

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

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

3 Change of scale

To calculate $E(X)$, $\text{Var}(X)$, $P(X \in dx)$ we sometimes make a change of scale

$Y = c + dX$ where c, d are constants

Y hopefully has a simpler density.

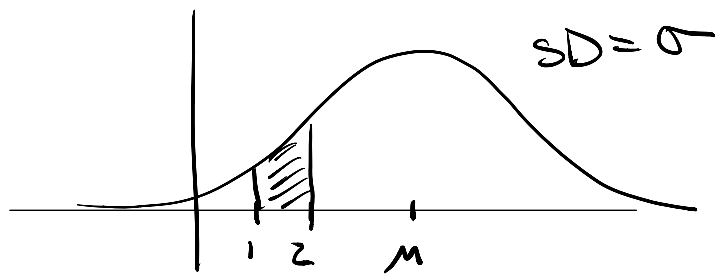
we can recover $E(X)$, $\text{Var}(X)$, $P(X \in dx)$ from $E(Y)$, $\text{Var}(Y)$, $P(Y \in dY)$.

ex

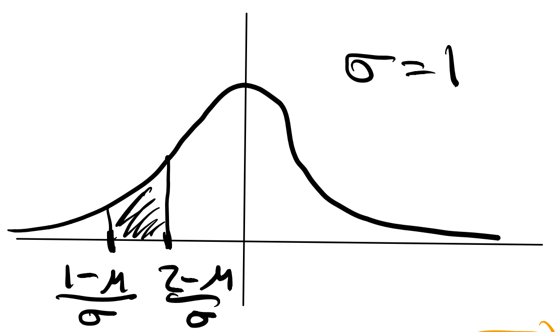
$$\text{Var}(Y) = \text{Var}(c + dX) = d^2\,\text{Var}(X)$$

$$\Rightarrow \text{Var}(X) = \frac{\text{Var}(Y)}{d^2}$$

ex Let $X \sim N(\mu, \sigma^2)$
Find $P(1 < X < 2)$



Lets make a change of scale so X is transformed to the std normal



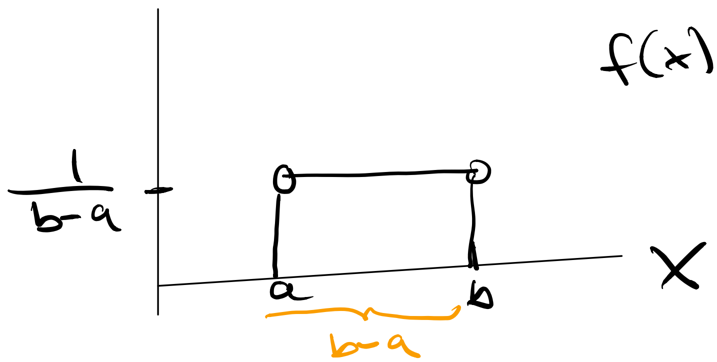
note
$$Var(Z) = Var\left(-\frac{\mu}{\sigma} + \frac{1}{\sigma}X\right) = \frac{1}{\sigma^2} Var(X) = \frac{\sigma^2}{\sigma^2} = 1$$

change of scale
$$Z = \frac{X-\mu}{\sigma} = \underbrace{-\frac{\mu}{\sigma}}_{c} + \underbrace{\frac{1}{\sigma}}_{d} X$$

$\rightsquigarrow$ $X = \mu + \sigma Z$

$$P(1 < X < 2) = P(1 < \mu + \sigma Z < 2) = P\left(\frac{1-\mu}{\sigma} < Z < \frac{2-\mu}{\sigma}\right) = \boxed{\Phi\left(\frac{2-\mu}{\sigma}\right) - \Phi\left(\frac{1-\mu}{\sigma}\right)}$$

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



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

a) You should change the scale of X to?

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

b) Find $E(X)$

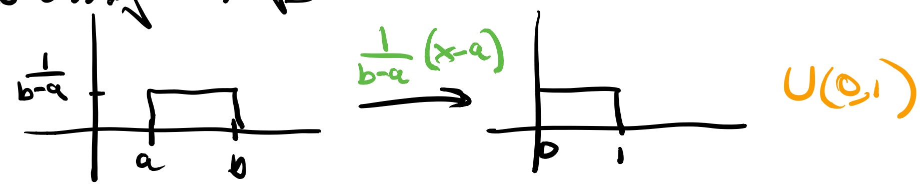
$$X = a + (b-a)U \Rightarrow E(X) = a + (b-a)\frac{1}{2}$$
$$= \frac{a+b}{2}$$

c) Find $Var(X)$.

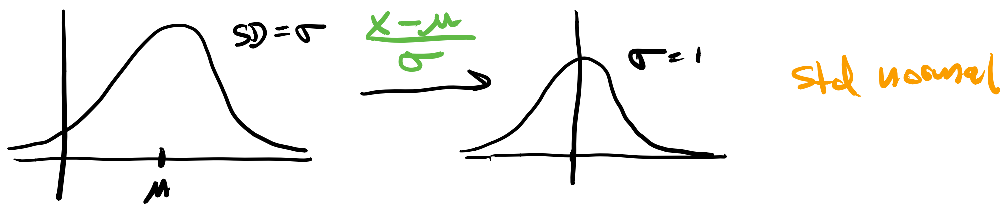
$$Var(X) = var((b-a)U + a)$$
$$= (b-a)^2 \underbrace{Var(U)}_{= 1/12} = \boxed{\frac{(b-a)^2}{12}}$$

<u>Idea</u> Make a change of scale so the density is as simple as possible so calculating things is easier

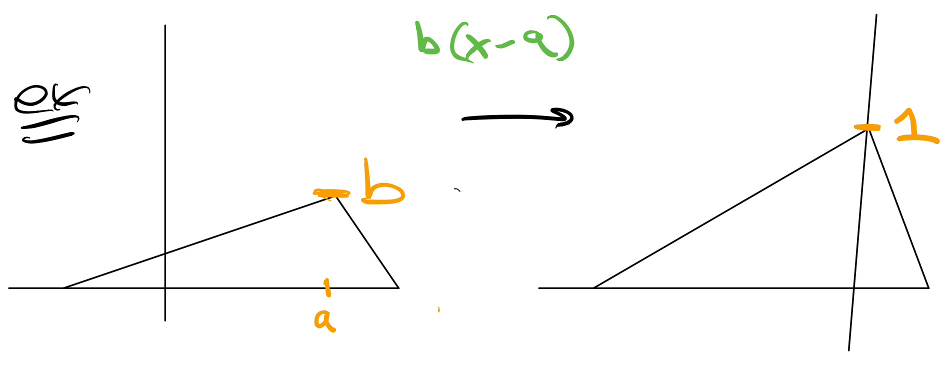
eg



eg



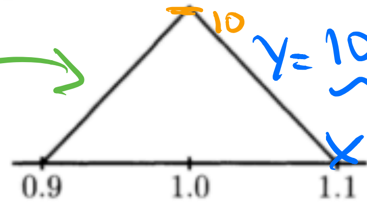
eg



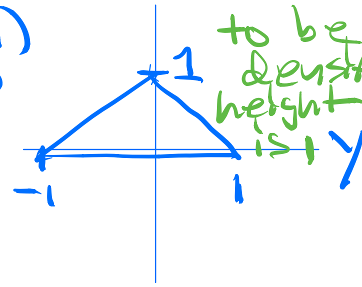
tinyurl.com/oct18-pt1

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.

to be density height is 10



$Y = 10(x-1)$



to be density height is 1

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

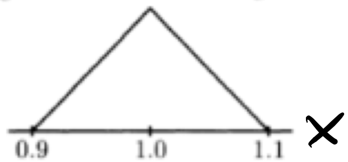
- a: X-1
- b: .1(X-1)
- c: 10X-1
- d: none of the above

More convenient to have height of triangle be 1 and center 0.

Our density is now

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

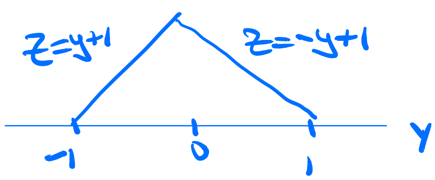
Suppose a manufacturing process designed to produce rods of length 1 inch exactly, in fact produces rods with length distributed according to the density graphed below.



Find the variance of the length of the rods.

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

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



Find Var(X)

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

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

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

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

ex A m