W8 Friday Snow Day

Stat 202-2015S-W9-Tues (Pg 1)

Review

The probability distribution of a discrete random variable is described by a table that lists both the possible values of X and their respective probabilities

Eg:

Value of X	0	1	2	3
Probability	1/8	3/8	3/8	1/8

X is the number of heads in three tosses of a coin.

~~The mean of a random variable~~

The probability distribution of a continuous random variable is described by its density curve

Discrete - Finite number of possible values
Continuous - an interval, possibley infinite, of possible values

If the sample space is discrete any random variable is discrete, if continuous, RV could be either continuous or discrete

The mean of a random variable

if discrete

$$\text{mean} = \sum x_i p_i$$

where the distribution is

Values of X	x_1	x_2	x_3	⋯	x_k
Probability	p_1	p_2	p_3	⋯	p_k

Example

Values	0	1	2	3
Probability	1/8	3/8	3/8	1/8

$$\text{mean} = 0 \cdot 1/8 + 1 \cdot 3/8 + 2 \cdot 3/8 + 3 \cdot 1/8$$
$$= 1.5$$

If the random variable is continuous you need calculus to compute its mean

In case you know calculus:

If $p(x)$ is the density

$$\text{mean} = \int_{-\infty}^{\infty} x\,p(x)\,dx$$

Law of Large Numbers
If the experiment is repeated many times and each time the value of X is observed

Then mean of X $\approx$ average of observations of X

Approximation gets better and better with more data

Big deal? Suppose you pick a person at random from a large population (random phenomenon) and observe their height. The mean of this random variable is the mean height of all people in the population (parameter)
Suppose there are too many people to observe all of them

So you pick n at random (sample)
Observe height of each

$$\text{estimate} = \overbrace{\frac{1}{n}\sum x_i}^{\text{average of observation}}$$

By the law of large numbers this gets closer and closer to the population mean with statistical certainty as $n \to \infty$

Rules for means

The values of random variables are numbers. We can add / multiply / subtract / etc these numbers

Therefore you can add/multiply subtract / etc random variables defined on the same sample space

$X+Y$ is a random variable for random variables X and Y
$aX+b$ is a RV for RV X and numbers a, b
$b_1X_1 + b_2X_2 + \dots + b_K X_K$ is a RV for RVs $X_1, \dots X_K$ and number $b_1 \dots b_K$.

The mean of a random variable X is written μ_X

Rule 1: $\mu_{X+Y} = \mu_X + \mu_Y$

Dice problem: Sum of numbers on a die is ~~a~~ RV.
mean is 3.5
mean of two dice is 7, even though this is a much more complicated RV

Gas

Pg5

(One dice has six possible values with equal probabilites)(two dice has 11 possible values with unequal probabilites)

Linear transformation

a and b are numbers
X is a RV

$$\mu_{a+bX} = a + b\mu_X$$

Difference

$$\mu_{X-Y} = \mu_X - \mu_Y$$

Linear Combo

$$\mu_{b_1X_1 + b_2X_2 + \cdots + b_nX_n}$$

$$= b_1\mu_1 + b_2\mu_2 + \cdots + b_n\mu_n$$

Variance of a RV – computed by

Discrete: computed from probability table
Continuous: Integral (calculus) involving density function

Discrete

Value of X	x_1	x_2	x_3	...	x_k
Probability	p_1	p_2	p_3	--	p_k

symbol for variance

$$\sigma_X^2 = (x_1 - \mu_X)^2 p_1 + (x_2 - \mu_X)^2 p_2 + \cdots + (x_k - \mu_X)^2 p_k$$

$$= \sum (x_i - \mu_X)^2 p_i$$

Example: toss 3 coins

Values of X	0	1	2	3
Probability	$1/8$	$3/8$	$3/8$	$1/8$

$$\mu_X = 1.5 = 3/2$$

$$\sigma_X^2 = \left.\begin{matrix}(0-3/2)^2 \cdot 1/8 \\ + (1-3/2)^2 \cdot 3/8 \\ + (2-3/2)^2 \cdot 3/8 \\ + (3-3/2)^2 \cdot 1/8\end{matrix}\right\} = \left\{\begin{matrix}9/4 \cdot 1/8 \\ + 1/4 \cdot 3/8 \\ + 1/4 \cdot 3/8 \\ + 9/4 \cdot 1/8\end{matrix}\right.$$

$$= \frac{1}{4} \cdot \frac{1}{8} \cdot (9+3+3+9) = \frac{1}{4} \cdot \frac{1}{8} \cdot 12 \cdot 2$$

$$= 3/4$$ ← mistake in last Wednesdays notes (will be corrected)

Standard Deviation $\sigma_X = \sqrt{\sigma_X^2} = \sqrt{3/4}$

Rules for Variances

~~Sum~~ <u>Linear Transformation</u>
(Any RV)

$$\sigma^2_{a+bX} = b^2\sigma^2_X$$

<u>Sum</u> (Independent RVs)

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$$

<u>Combo</u> (independent RVs)

$$\sigma^2_{b_1X+b_2Y} = b_1^2\sigma^2_X + b_2^2\sigma^2_Y$$

<u>Difference</u> (independent RV) (type of Combo)

$$\sigma^2_{X-Y} = (1)^2\sigma^2_X + (-1)^2\sigma^2_Y$$

$$= \sigma^2_X + \sigma^2_Y$$

not a mistake

Why are these formulas useful?

Well in Chapter 5 ~~...~~ (and ~~...~~ elsewhere)
we are going to need the mean
and variance of a Binomial
Random variable
remember this was like tossing an
~~...~~ where the probability unfair
of "success" (heads) was p coin
and probability of "failure" (tails) was 1-p, many times
(Remember Binomial calculator from exam.)
The Binomial RV is the count of # of successes.
Well I could just give you the formulas
for mean and variance
But they would be very bizzare
It wouldn't make sense why they
were true

With the formulas shown today
I can actually explain why
these formulas are true.

~~#~~ They will be similarly helpful in
other situations.

Lets derive the mean and variance of the random variable "number of heads in three tosses of a coin" that we have seen earlier

We have already derived them

$\mu_X = 3/2 \qquad \sigma_X^2 = 3/4$ But it was hard and our method will get harder with large values of n

Now we will use a method that works easily for large n,

Consider that X is the sum of three independent RVs each one the number of heads in the toss of one coin,

$$X = Y_1 + Y_2 + Y_3$$

Values of Y_i	0	1
probability	$\frac{1}{2}$	$\frac{1}{2}$

Mean μ_{Y_1} $0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$

$$\sigma_{Y_i}^2 = \left(0-\tfrac{1}{2}\right)^2 \cdot \tfrac{1}{2} + \left(1-\tfrac{1}{2}\right)^2 \cdot \tfrac{1}{2}$$

$$= \frac{1}{4}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{2} = \frac{1}{4}$$

Now we apply our rules

$$\mu_X = \mu_{Y_1} + \mu_{Y_2} + \mu_{Y_3} = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2}$$

check ✓

Because independent Y_1

$$\sigma_X^2 = \sigma_{Y_1}^2 + \sigma_{Y_2}^2 + \sigma_{Y_3}^2 = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$

check ✓

Now in general

Values of Y_i	0	1
Probability	$1-p$	p

$$\mu_{Y_1} = 0 \cdot (1-p) + 1 \cdot p = p$$

$$\sigma_{Y_1}^2 = (0-p)^2(1-p) + (1-p)^2 \cdot p$$

$$= p^2(1-p) + (1-p)^2 \cdot p$$

$$= p(1-p)[p + 1 - p]$$

$$= p(1-p)$$

$$\mu_X = \mu_{Y_1} + \dots + \mu_{Y_n} = np$$

Because Independent

$$\sigma_X^2 = \sigma_{Y_1}^2 + \dots + \sigma_{Y_n}^2 = np(1-p)$$

$$\sigma_X = \sqrt{np(1-p)}$$

Last thing today:

We had ~~a~~ rules for the variance of a sum and difference of two random variables that were independent:

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$$

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y$$

Lets talk about the difference formula a little more

Suppose X and Y are height of two people in a population and they are independent.

If we had "$\sigma^2_{X-Y} = \sigma^2_X - \sigma^2_Y$"

Then ~~the~~ we would conclude the variance is zero. This would be wrong because if you take two ~~people~~ numbers at random and take their difference, you won't get the same number each time. There is spread in the distribution so variance is nonzero $\sigma^2_{X-Y} \neq \sigma^2_X - \sigma^2_Y$

Now consider a different experiment where X is the height of a random person who has an identical twin and Y is the height of his or her twin. (These are not independent)

Now suppose there are no environmental factors affecting height so always $X=Y$

Then $X-Y=0$ always so

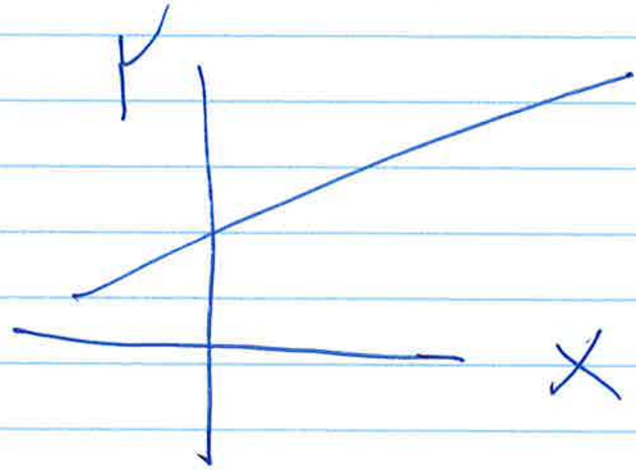
$$\sigma^2_{X-Y}=0$$

In this case $\sigma^2_{X-Y}=\sigma^2_X-\sigma^2_Y$ so the formula $\sigma^2_{X-Y}=\sigma^2_X+\sigma^2_Y$ evidently doesn't work for non independent RVs.

The general formula involves the correlation between X and Y called ρ

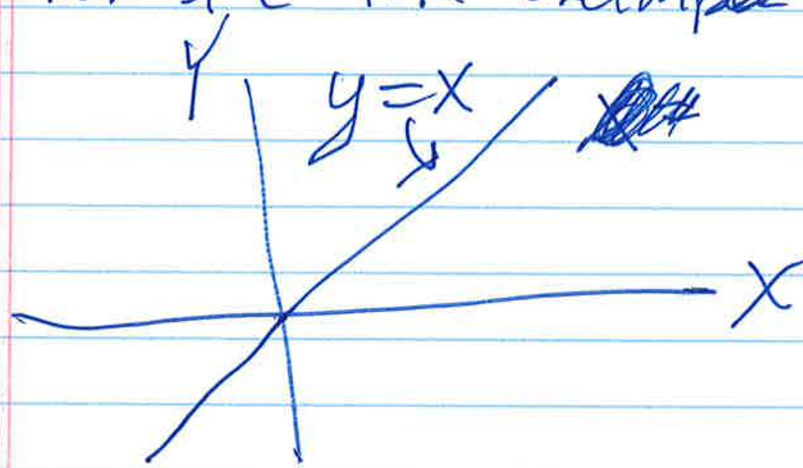
Like the correlation in data, ρ is a number representing the strength of the linear relationship between X and Y, specifically it is a number between -1 and $+1$

ρ is +1 if the relationship between x and Y is deterministic (not statistical) and is a line with positive slope

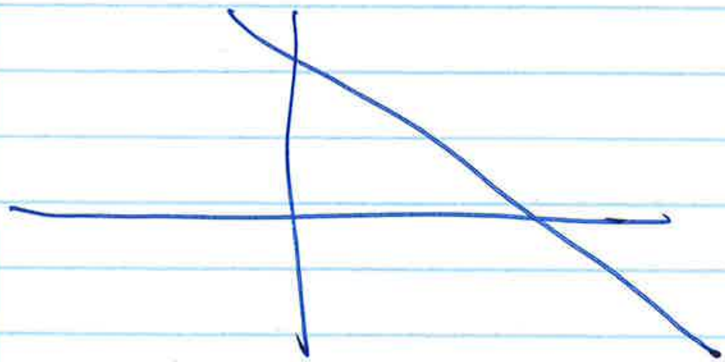


There can be spread in x and Y separately but if you plot them together they will fall on a line

For the twin example $\rho = 1$



ρ is -1 if the relationship between X and Y is deterministic and is a line with negative slope

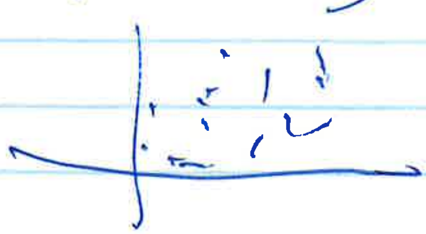


ρ is 0 if there is no linear relationship between X and Y

In particular ρ is 0 if X and Y are independent

But knowing $\rho = 0$ is not enough to conclude X and Y are independent

$|\rho| < 1$ if there is scatter in X and Y



The general rules for the variance of a sum and difference of two RVs that are not necessarily independent are given by

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y + 2\rho\sigma_X\sigma_Y$$

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y - 2\rho\sigma_X\sigma_Y$$

For the twin case $\rho = 1$

$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y - 2\sigma_X\sigma_Y$$
$$= (\sigma_X - \sigma_Y)^2$$

Note it is not always the case when $\rho = 1$ that $\sigma^2_{X-Y} = 0$

However when $\rho = 1$ and $\sigma_X = \sigma_Y$ then

$$\sigma^2_{X-Y} = 0.$$