Stat 134 lec 34

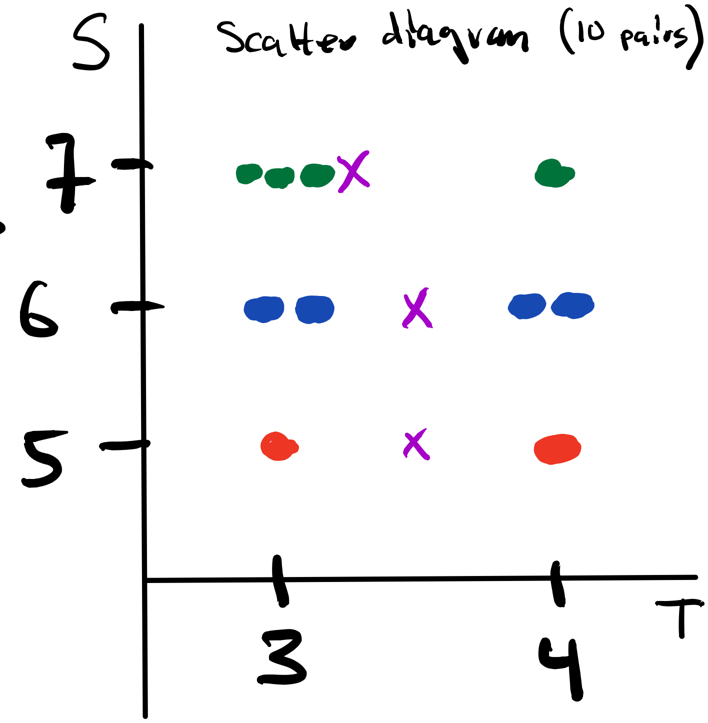
Sec 6.2 Conditional Expectation

joint distribution

	T=3	T=4	Sum
S=7	0.3	0.1	0.4
S=6	0.2	0.2	0.4
S=5	0.1	0.1	0.2
Sum	0.6	0.4	1

$\longleftrightarrow$

Scatter diagram (10 pairs)



Calculate

a) $E(T)$

$$E(T) = 3\cdot\frac{6}{10} + 4\cdot\frac{4}{10} = 3.4$$

b) $E(T|S=6)$

$$E(T|S=6) = 3\cdot\frac{\frac{2}{10}}{\frac{4}{10}} + 4\cdot\frac{\frac{2}{10}}{\frac{4}{10}} = 3.5$$

$g(S) = E(T|S)$

$g(7) = E(T|S=7) = 3.25$

$g(6) = 3.5$

$g(5) = 3.5$

$$E(g(S)) = \sum_{s \in S} g(s)\cdot P(S=s)$$

$$= 3.25(.4) + 3.5(.4) + 3.5(.2)$$

$$= 3.4$$ ← this is $E[T]$.

it is the weighted average of all of the group averages.

In other words,

$$E(E(T \mid S)) = E(T)$$

This is called the property of <u>iterated expectations</u>.

Intuitively,

If you have a class that is $2/3$ girls and $1/3$ boys and the girls weigh on average 100 lbs and boys weigh 200 lbs then the average weight of the class should be $\frac{2}{3}(100) + \frac{1}{3}(200)$. <u>i.e.</u> we take the weighted average of the averages.

Here is a formal proof: (Please read)

Iterated Expectation

$$E(Y) = E(E(Y|X))$$

$$E(Y) = \sum_{\text{all } y} y P(Y=y)$$

$$= \sum_{\text{all } y} y \sum_{\text{all } x} P(X=x, Y=y)$$

$$= \sum_{\text{all } y} y \sum_{\text{all } x} \frac{P(X=x, Y=y)}{P(X=x)} P(X=x)$$

$$= \sum_{\text{all } y} y \sum_{\text{all } x} P(Y=y | X=x) P(X=x)$$

$$= \sum_{\text{all } x} \left(\sum_{\text{all } y} y P(Y=y | X=x) \right) \cdot P(X=x)$$

$$= \sum_{\text{all } x} E(Y | X=x) \cdot P(X=x)$$

$$= E(E(Y|X))$$

ex $N \sim \text{Pois}(\mu)$

Given $N=n$, X is the number of heads in n tosses a p-coin.

Find $E(X)$.

What is the distribution of $X|N=n$? — $\text{bin}(n,p)$

$$E(X|N=n) = np$$

$$g(N) = E(X|N) = Np$$

$$E(X) = E(g(N)) = E(Np) = pE(N)$$

$$E(N) = \mu.$$

$$= \boxed{p\mu}$$

Stat 134

Friday April 13 2018

1. 8 transistors (type 1) are distributed Exp(1/100) and 4 transistors (type 2) are Exp(1/200). Let T be the lifetime of a randomly picked transistor. E(T) is:

   a 120

   b 133.3

   c 150

   d none of the above

$E(T) = \frac{2}{3}(100) + \frac{1}{3}(200)$

Properties of conditional expectation

For a fixed conditioning event A, conditional expectation has all the familiar properties of expectation

$$E(X+Y)=E(X)+E(Y)$$

$$E(X+Y\mid A)=E(X\mid A)+E(Y\mid A)$$

since $X+Y\mid A$ is a restriction of $X+Y$ to a smaller outcome space.

So suppose A is the event $X=x$

what is $E(X+Y\mid X=x)$?

$$= E(X\mid X=x)+E(Y\mid X=x)$$

$$= x+E(Y\mid X=x)$$

so we write,

equality of RV.

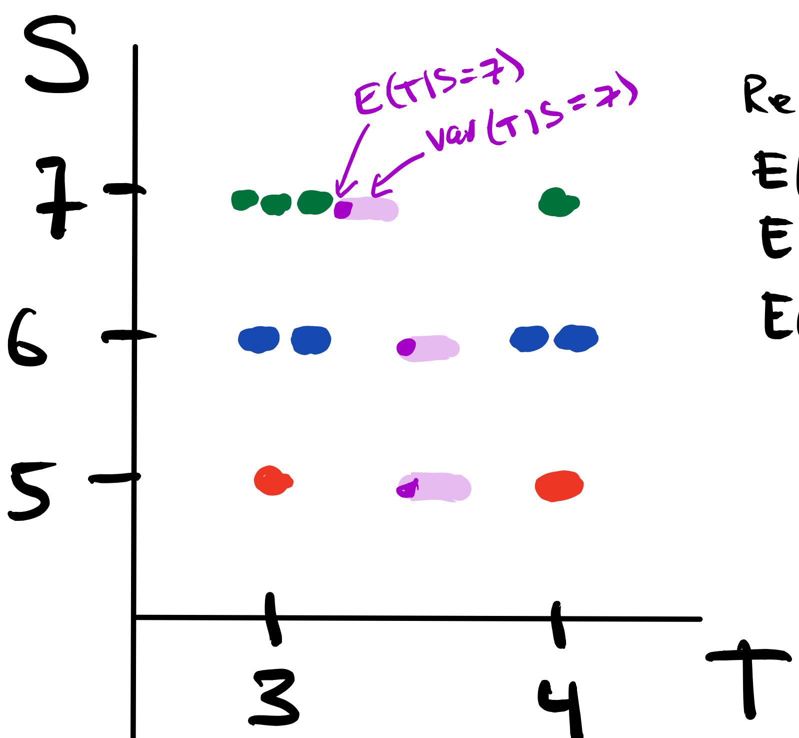
$$E(X+Y\mid X)=X+E[Y\mid X]$$

Properties

		Equality of
1	$E(X) = E(E(X\mid Y))$	numbers
2	$E(aY+b\mid X) = aE(Y\mid X)+b$	RVs
3	$E(Y+Z\mid X) = E(Y\mid X)+E(Z\mid X)$	RVs
4	$E(g(X)\mid X) = g(X)$	RVs
5	$E(g(X)Y\mid X) = g(X)E(Y\mid X)$	RVs
6	$\mathrm{Var}(X) = E(\mathrm{Var}(X\mid Y)) + \mathrm{Var}(E(X\mid Y))$	RVs

total variance decomposition (see 6.2.18)

see below



Recall

$E(T\mid S=7) = 3.25$

$E(T\mid S=6) = 3.5$

$E(T\mid S=5) = 3.5$

Calculate Var(T)

Generally you would want to use the formula $Var(T) = E(T^2) - E(T)^2$ but here T is a Bernoulli RV (i.e it takes 2 values) and

$$Var(T) = (\text{big} - \text{small})^2 \cdot P(\text{big})P(\text{small})$$
$$= (4-3)^2\left(\frac{4}{10}\right)\left(\frac{6}{10}\right) = .24$$

Find $Var(T \mid S=7)$

$$= \left(\frac{3}{4}\right)\left(\frac{1}{4}\right) = \frac{3}{16} = .1875.$$

Half class

Find $Var(E(T|S))$

$$E(E(T|S)^2) =$$
$$(3.25)^2(.4) + (3.5)^2(.4) + (3.5)^2(.2)$$
$$= 11.575$$
$$E(E(T|S))^2 = E(T)^2 = (3.4)^2$$
$$= .015$$

Half class

Find $E(Var(T|S))$

$$Var(T|S=7) = .1875$$
$$Var(T|S=6) = \left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = .25$$
$$Var(T|S=4) = \left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = .25$$
$$E(Var(T|S)) = .1875(.4) + (.25)(.4) + (.25)(.2)$$
$$= .225$$

Notice $.015 + .225 = .24 = \overline{\text{var}(T)}$

$\underset{\text{var}(E(T|S))}{\underset{\shortparallel}{.015}} \quad \underset{E(\text{var}(T|S))}{\underset{\shortparallel}{.225}}$

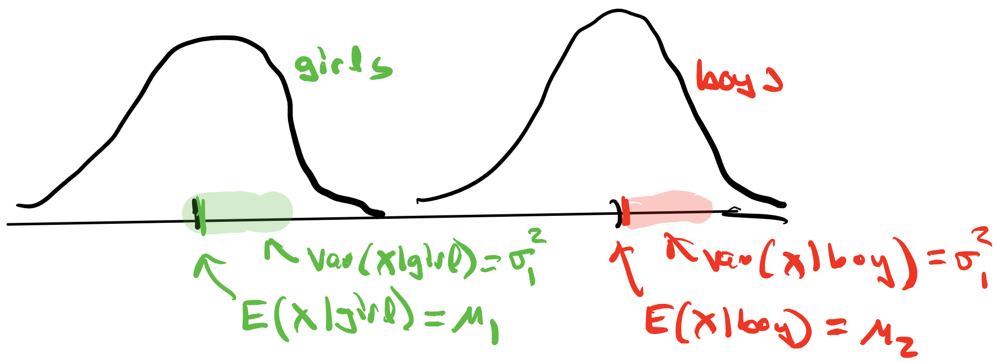
This gives us the total variance decomposition:

$$\boxed{\text{Var}(T) = \text{Var}(E(T|S)) + E(\text{var}(T|S))}$$

$\text{Var}(E(T|S))$ is the between group variance. It is the variance of 3.25, 3.5, 3.5 (i.e the variance of our group averages).

$E(\text{Var}(T|S))$ is the within group variance. It is the average of .1875, .25, .25 (i.e the average of our grp variances).

The decomposition of the total variance into between group variance and within group variance is the key idea to Analysis of Variance (ANOVA) in statistics!! (covered in stat 135) ex Imagine the weight distribution of students in our class looks like



$Var(X)$ = variance within groups

$\qquad$ = average of $\{\sigma_1^2, \sigma_2^2\}$

$\quad$ + variance between groups

$\qquad$ = variance of $\{\mu_1, \mu_2\}$

You will prove this in 6.2.18 !