M 384G/374G/CAM384T

CONDITIONAL AND MARGINAL MEANS AND VARIANCES

I. Question: How are conditional means E(Y|X) and marginal means E(Y) related?

Simple example:

Population consisting of n_1 men, n_2 women.
Y = height
X = sex
Categorical, two values: Male, Female

So there are two conditional means:
E(Y|male) = (Sum of all men's heights)/n_1
E(Y|female) = (Sum of all women's heights)/n_2

Then
Sum of all men's heights = n_1E(Y|male)
Sum of all women's heights = n_2 E(Y| female)

The marginal mean is
E(Y) = (Sum of all heights)/(n_1 + n_2) =

$$\frac{(Sum\ of\ men's\ heights) + (Sum\ of\ women's\ heights)}{n_1 + n_2}$$

$$= \frac{n_1 E(Y \mid male) + n_2 E(Y \mid female)}{n_1 + n_2}$$

$$= \frac{n_1}{n_1 + n_2}\text{E(Y| male)} + \frac{n_2}{n_1 + n_2}\text{E(Y| female)}$$

= (proportion of males)(E(Y| male) + (proportion of females)(E(Y| female)

= (probability of male)(E(Y| male) + (probability of female)(E(Y| female)

Thus: *The marginal mean is the weighted average of the conditional means, with weights equal to the probability of being in the subgroup determined by the corresponding value of the conditioning variable.*

Similar calculations show: If we have a population made up of m subpopulations pop_1, pop_2, …, pop_m (equivalently, if we are conditioning on a categorical variable with m values -- e.g., the age of a fish), then

$$\text{E(Y)} = \sum_{k=1}^{m} \Pr(pop_k) E(Y \mid pop_k)$$

e.g., for our fish, pop_k = fish of age k, and

$$\text{E(length)} = \sum_{k=1}^{6} \Pr(Age = k) E(Length \mid Age = k)$$

Rephrasing in terms of the categorical variable X defining the subpopulations,

$$\text{E(Y)} = \sum_{all\ values\ x\ of\ X} \Pr(x) E(Y \mid X = x)$$

Stated in words:

The *analogue for conditioning on a continuous variable X* is:

$$\text{E(Y)} = \int_{-\infty}^{\infty} f_X(x) E(Y \mid x)\, dx,$$

where $f_X(x)$ is the probability density function (pdf) of X.

Note:
1. There are analogous results for conditioning on more than one variable.

2. The analogous result for *sample* means is

$$\bar{y} =$$

II. A second (related) relationship between marginal and conditional means for populations:

Consider E(Y|X) as a new random variable U as follows:
Randomly pick an x from the distribution of X.
The new r.v. U has value E(Y|X = x).

Example: Y = height, X = sex
Randomly pick a person from the population in question.

$$U = \begin{cases} \mu_f = E(Y \mid X = female)\ if\ the\ person\ is\ female \\ \mu_m = E(Y \mid X = male)\ if\ the\ person\ is\ male \end{cases}$$

Question to think about: What might cause P(U = u) to be high? [Hint: There are two ways this might arise]

Consider the expected value of this new random variable. (e.g., the expected value of the mean height for the sex of a randomly selected person from the given population. In this case, we would expect E(U) to depend on the proportion of the population which is of each sex.)

If U is discrete, then

$$E(U) = \sum_{\substack{\text{All possible} \\ \text{values of } U}} P(u)u \qquad \text{[Why?]}$$

Example: For U = E(height | sex), the values taken on by U are

_______________ and ________________,

with respective probabilities ___________ and ______________,

so E(U) = ____,

which from Part I is just _______________________________ .

In other words,

E(E(ht|sex) =

The same reasoning works in general, showing that:

The expected value of the conditional means is the weighted average of the conditional means, which from Part 1 is just the marginal mean
i.e.,

E(E(Y|X)) = weighted average of conditional means
= E(Y)

III. CONDITIONAL AND MARGINAL VARIANCE

Marginal Variance: The *definition* of the (population) (marginal) variance of a random variable Y is

$\text{Var}(Y) = E([Y - E(Y)]^2)$

What does this say in words (and pictures)?

There is another *formula* for Var(Y) that is sometimes useful in computing variances or proving things about them. It can be obtained by multiplying out the squared expression in the definition:

$\text{Var}(Y) = E([Y - E(Y)]^2) = E(Y^2 - 2YE(Y) + [E(Y)]^2)$

= ____

(Fill in details, and say the final result in words!)

Conditional Variance: Similarly, if we are considering a conditional distribution Y|X, we define the *conditional variance*

$$\mathrm{Var}(Y|X) = \text{Variance of the conditional distribution } Y|X$$

$$= E([Y - E(Y|X)]^2 \mid X)$$

(Note that *both* expected values here are conditional expected values.)

What does this say in words (and pictures)?

Exercise: Derive another formula for the conditional variance, analogous to the second formula for the marginal variance. (And say it in words!)

Conditional Variance as a Random Variable: As with E(Y|X), we can consider Var(Y|X) as a random variable. For example, if Y = height and X = sex for persons in a certain population, then Var(height | sex) is the variable which assigns to each person in the population the variance of height for that person's sex.

IV. CONNECTING MEANS AND VARIANCES

Expected Value of the Conditional Variance: Since Var(Y|X) is a random variable, we can talk about its expected value. Using the formula $\mathrm{Var}(Y|X) = E(Y^2|X) - [E(Y|X)]^2$, we have

$$E(\mathrm{Var}(Y|X)) = E(E(Y^2|X)) - E([E(Y|X)]^2)$$

We have already seen that the expected value of the conditional expectation of a random variable is the expected value of the original random variable, so applying this to Y^2 gives

$$(*) \qquad E(\mathrm{Var}(Y|X)) = E(Y^2) - E([E(Y|X)]^2)$$

Variance of the Conditional Expected Value: For what comes next, we will need to consider the variance of the conditional expected value. Using the second formula for variance, we have

$$\mathrm{Var}(E(Y|X)) = E([E(Y|X)]^2) - [E(E(Y|X))]^2$$

Since E(E(Y|X)) = E(Y), this gives

$$(**)\mathrm{Var}(E(Y|X)) = E([E(Y|X)]^2) - [E(Y)]^2.$$

Putting It Together:

Note that (*) and (**) both contain the term $E([E(Y|X)]^2)$, but with opposite signs. So adding them gives:

$$E(Var(Y|X)) + Var(E(Y|X)) = E(Y^2) - [E(Y)]^2,$$

which is just Var(Y). In other words,

$$(***) \quad Var(Y) = E(Var(Y|X)) + Var(E(Y|X)).$$

In words: The marginal variance is the sum of the expected value of the conditional variance and the variance of the conditional means.

Consequences:

1) This says that two things contribute to the marginal (overall) variance: the expected value of the conditional variance, and the variance of the conditional means. (See Exercise) Moreover, $Var(Y) = E(Var(Y|X))$ if and only if $Var(E(Y|X)) = 0$. What would this say about $E(Y|X)$?

2) Since variances are always non-negative, (***) implies

$$Var(Y) \geq E(Var(Y|X)).$$

3) Since $Var(Y|X) \geq 0$, $E(Var(Y|X))$ must also be ≥ 0. (Why?). Thus (***) implies

$$Var(Y) \geq Var(E(Y|X)).$$

Moreover, $Var(Y) = Var(E(Y|X))$ if and only if $E(Var(Y|X)) = 0$. What would this imply about $Var(Y|X)$ and about the relationship between Y and X?

4) Another perspective on (***) (cf. Textbook, pp. 36 - 37):

Note that:

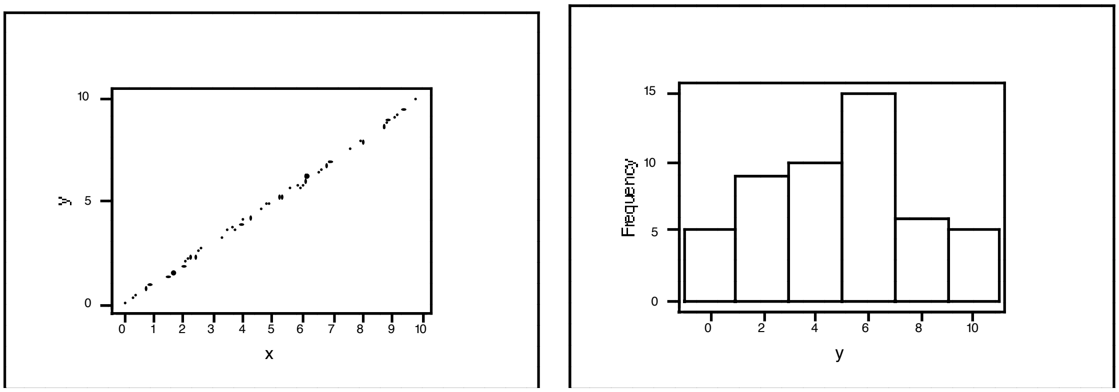
i) $E(Var(Y|X)$ is a weighted average of $Var(Y|X)$

ii) $Var(E(Y|X) = E([E(Y|X) - E(E(Y|X))]^2)$
$= E([E(Y|X) - (E(Y)]^2)$,
which is a weighted average of $[E(Y|X) - (E(Y)]^2$

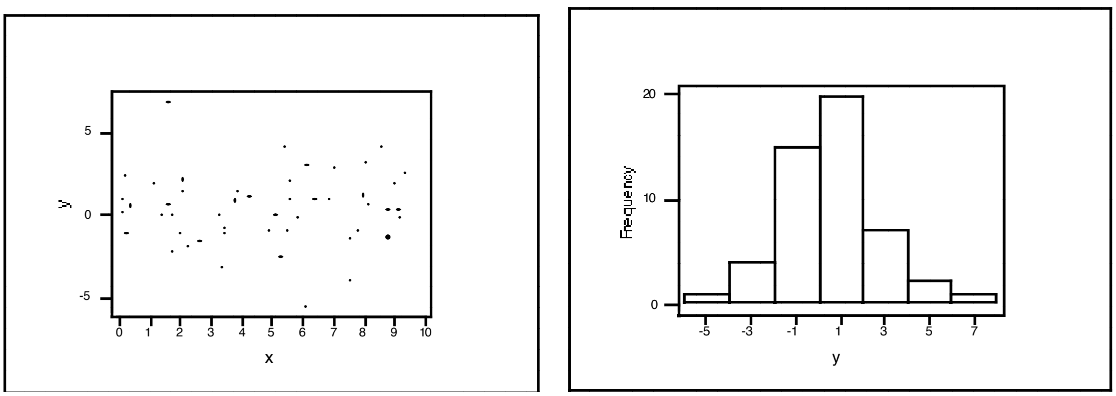
Thus, (***) says that $Var(Y)$ is a weighted mean of $Var(Y|X)$ plus a weighted mean of $[E(Y|X) - (E(Y)]^2$ (and is a weighted mean of $Var(Y|X)$ if and only if all conditional expected values $E(Y|X)$ are equal to the marginal expected value $E(Y)$.)

EXERCISE: Which contributes most to Var(Y): Var(E(Y|X)) or E(Var(Y|X))?

A.



B.



C.