4. Expectation and variance

4.1. Expectation in general

For a discrete random variable we defined

$$\boxed{\begin{aligned} E(X) &= \sum_{x_k} x_k P(X = x_k) \\ E[f(X)] &= \sum_{x_k} f(x_k) P(X = x_k) \end{aligned}}$$

For a discrete random vector we have

$$\boxed{E[f(\underline{X})] = \sum_{\underline{x}_k} f(\underline{x}_k) P(\underline{X} = \underline{x}_k)}$$

We need to extend the notion of expectation to continuous random variables and vectors.

Definitions:

(i) Let X have density $f_X(x)$.

We define

$$\boxed{\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x\, f_X(x)\, dx \\ E[f(X)] &= \int_{-\infty}^{\infty} f(x)\, f_X(x)\, dx \end{aligned}}$$

Technical note: we say that $E(X)$ exists if the integral $\int_{-\infty}^{\infty} |x|\, f_X(x)\, dx$ converges and similarly for $E[f(X)]$.

(ii) Let the random vector $\underline{X}$ have density $f_{\underline{X}}(\underline{x})$. We define

$$\boxed{E[f(\underline{X})] = \int_{\mathbb{R}^r} f(\underline{x})\, f_{\underline{X}}(\underline{x})\, d\underline{x}}$$

Examples :

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

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

$$= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

$= (*)$ New variable: $\frac{x-\mu}{\sigma} = u$

$$= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} (\sigma u + \mu)\, e^{-u^2/2}\, du$$

$$= \mu \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2}\, du$$

$$= \mu$$

because $\int_{-\infty}^{\infty} u \cdot e^{-u^2/2}\, du = 0$

(odd function).

We continue

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

$$= \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \int_{-\infty}^{\infty} x^2 \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma u+\mu)^2\, e^{-u^2/2}\,du$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma^2 u^2 + \mu^2) \cdot e^{-u^2/2}\,du$$

(the middle term = 0)

$$= \sigma^2 \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 \cdot e^{-u^2/2}\,du + \mu^2 = (*)$$

We integrate by parts

$$\int_{-\infty}^{\infty} u^2 \cdot e^{-u^2/2}\,du =$$

$$= \int_{-\infty}^{\infty} u \cdot u \cdot e^{-u^2/2}\,du$$

$$= -u \cdot e^{-u^2/2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-u^2/2}\,du$$

$$= \sqrt{2\pi}$$

$$(*) = \sigma^2 + \mu^2$$

Definition: Let X be a random variable. Let $\mu = E(X)$.
We call $E(X^m)$ the m-th moment of X and $E[(X-\mu)^m]$ the m-th central moment of X.

Example: Let $X \sim \Gamma(a, \lambda)$.
We compute the m-th moment of X as

$$E(X^m) = \int_0^\infty x^m \cdot f_X(x)\, dx$$

$$= \int_0^\infty \frac{\lambda^a}{\Gamma(a)} \cdot x^m \cdot x^{a-1} \cdot e^{-\lambda x}\, dx$$

$$= \frac{\lambda^a}{\Gamma(a)} \int_0^\infty x^{m+a-1} e^{-\lambda x}\, dx$$

$$= \frac{\lambda^a}{\Gamma(a)} \cdot \frac{\Gamma(m+a)}{\lambda^{m+a}}$$

$$= \frac{\Gamma(m+a)}{\Gamma(a)\,\lambda^{m}}$$

$$= \frac{(a)_m}{\lambda^m}$$

where $(a)_m = a(a+1)\cdots(a+m-1)$.

Example: Let (X,Y) have the density $f_{X,Y}(x,y) =$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}}\, e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}$$

with $|\rho|<1$. We have

$$E(XY) = \int_{\mathbb{R}^2} xy \cdot f_{X,Y}(x,y)\,dx\,dy$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{\mathbb{R}^2} xy \cdot e^{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}}\,dx\,dy$$

Fubini

$$= \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} x \cdot e^{-x^2/2}\,dx \;\times$$

$$\times \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} y \cdot e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}}\,dy$$

The last integral is ρx. We computed it in the first example. We have

$$E(XY) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \rho x^2 \cdot e^{-x^2/2} \, dx$$

$$= \rho \cdot E(X^2)$$

$$= \rho$$

If $X \sim N(0,1)$ then $E(X^2) = \sigma^2 + \mu^2 = \sigma^2$.

<u>Theorem 4.1</u>: Let (X,Y) have density $f_{X,Y}(x,y)$. We have

$$E[\alpha X + \beta Y] = \alpha E(X) + \beta E(Y)$$

<u>Proof</u>: We have

$E[\alpha X + \beta Y]$

$$= \int_{\mathbb{R}^2} (\alpha x + \beta y)\, f_{X,Y}(x,y)\, dx\, dy$$

$$= \alpha \cdot \int_{\mathbb{R}^2} x\, f_{X,Y}(x,y)\, dx\, dy$$

$$+ \beta \int_{\mathbb{R}^2} y\, f_{X,Y}(x,y)\, dx\, dy$$

$$= \alpha \cdot E(X) + \beta E(Y)$$

Remarks:

(i) The same proof works with $\alpha f(X,Y) + \beta g(X,Y)$ i.e

$$E[\alpha f(X,Y) + \beta g(X,Y)]$$

$$= \alpha\, E[f(X,Y)] + \beta\, E[g(X,Y)].$$

(ii) The theorem has a vector version:

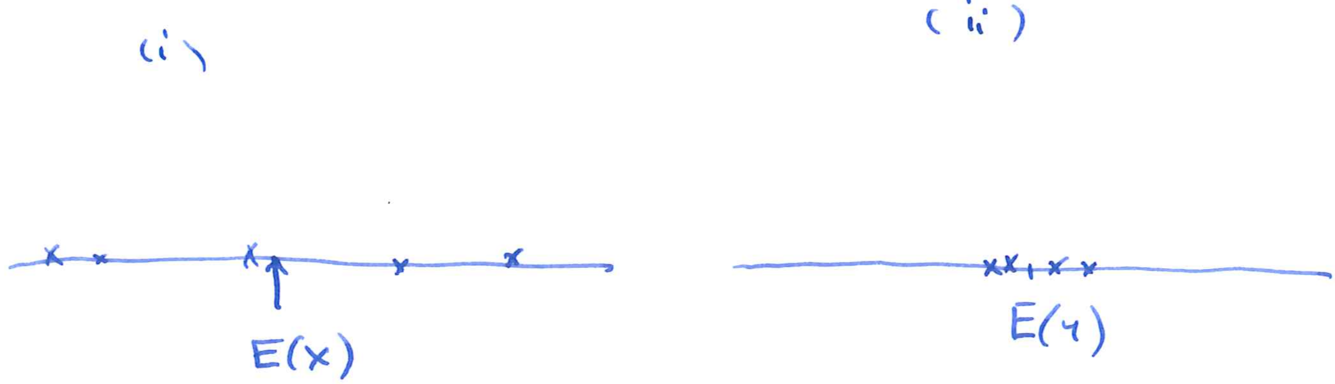
$$\boxed{E\left[\sum_{k=1}^{r} \alpha_k X_k\right] = \sum_{k=1}^{r} \alpha_k E(X_k).}$$

The expectation is always linear.

4.2 Variance and covariance

We motivated the expectation as a „long term average". We would like to devise a measure of dispersion of a random variable X.

Figure:



In the figure we have „repeated" values of X and Y. The values of X are more „dispersed". Why do we say this? On average the values of X are further away from $E(x)$. Repetitions to the right and to the left contribute an equal amount to dispersion so we take absolute distances. However, Gauss chose the square. His choice was motivated by mathematical considerations. Denote $v_1, v_2, \dots, v_n$ the repetitions of X.

The dispersion according to Gauss is

$$\frac{(v_1 - E(x))^2 + \cdots + (v_n - E(x_n))^2}{n}.$$

If we take $f(x) = (x - E(x))^2$ we see that the above average

$$\approx E[(x - E(x))^2].$$

<u>Definition</u>: The variance of the random variable X is given by

$$\boxed{\operatorname{var}(X) = E[(X - E(X))^2]}.$$

We compute

$$E[(X - E(X))^2] =$$

$$= E[X^2 - 2E(X) \cdot X + E(X)^2]$$

lin.

$$= E(x^2) - 2E(x)\cdot E(x) + E(x)^2$$

$$= E(x^2) - [E(x)]^2$$

<u>Alternative form</u>:

$$\boxed{\operatorname{var}(x) = E(x^2) - [E(x)]^2}$$

<u>Examples</u>:

(i) $x \sim \mathrm{Bin}(n, p)$

We know

$E(x^2) = npq + n^2p^2$ and

$E(x) = n\cdot p$

$$\operatorname{var}(x) = E(x^2) - E(x)^2$$

$$= npq$$

(ii) $X \sim NegBin(m, p)$

We know

$$E(X) = \frac{m}{p}$$

$$E(X^2) = \frac{m \cdot q}{p^2} + \frac{m^2}{p^2}$$

$$var(X) = \frac{m \cdot q}{p^2}$$

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

We know

$$E(X) = \mu$$

$$E(X^2) = \sigma^2 + \mu^2$$

$$var(X) = E(X^2) - [E(X)]^2$$

$$= \sigma^2$$

Comment: If $X \sim N(\mu, \sigma^2)$ the two parameters have a nice interpretation. They are the expectation and the variance.

What about the variance of sums? We compute

$$\operatorname{var}(X+Y) = E[(X+Y)^2] - [E(X+Y)]^2$$

$$= E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2$$

$$\overset{\text{lin.}}{=} E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2$$

$$= \operatorname{var}(X) + \operatorname{var}(Y) + 2[E(XY) - E(X)E(Y)]$$

There is no reason for the term in square brackets to be 0.

Definition: Let X, Y be random variables. The quantity

$$E(XY) - E(X) \cdot E(Y)$$

is called the covariance of X and Y and denoted by $\mathrm{cov}(X,Y)$.

Remark: An application of linearity gives that

$$\mathrm{cov}(X,Y) = E\left[(X-E(X))(Y-E(Y))\right].$$

Theorem 4.2: Let $X_1 \dots X_r$ and $Y_1, Y_2, \dots, Y_s$ be random variables. We have

$$\mathrm{cov}\left(\sum_{k=1}^{r} \alpha_k X_k, \sum_{\ell=1}^{s} \beta_\ell Y_\ell\right) = \sum_{k=1}^{r} \sum_{\ell=1}^{s} \alpha_k \beta_\ell \, \mathrm{cov}(X_k, Y_\ell)$$

Proof: We compute

$$E\Big[\Big(\sum_{k=1}^{r} \alpha_k X_k\Big)\Big(\sum_{\ell=1}^{s} \beta_\ell Y_\ell\Big)\Big]$$

$$= E\Big[\sum_{k=1}^{r}\sum_{\ell=1}^{s} \alpha_k \beta_\ell X_k Y_\ell\Big]$$

$$\overset{\text{lin.}}{=} \sum_{k=1}^{r}\sum_{\ell=1}^{s} \alpha_k \beta_\ell E(X_k Y_\ell)$$

On the other hand

$$E\Big(\sum_{k=1}^{r} \alpha_k X_k\Big) \cdot E\Big(\sum_{\ell=1}^{s} \beta_\ell Y_\ell\Big)$$

$$\overset{\text{lin}}{=} \sum_{k=1}^{r}\sum_{\ell=1}^{s} \alpha_k \beta_\ell E(X_k) E(Y_\ell)$$

We subtract and get the result.

Remark: The property is called bilinearity.

The definitions give us further properties of covariances that follow from definitions:

$$\boxed{\begin{array}{lll} \text{(i)} & \operatorname{var}(aX) = a^2 \operatorname{var}(X) \\ \text{(ii)} & \operatorname{cov}(X,X) = \operatorname{var}(X) \\ \text{(iii)} & \operatorname{cov}(X,Y) = \operatorname{cov}(Y,X) \\ \text{(iv)} & \operatorname{cov}(\alpha X, \beta Y) = \alpha\beta \operatorname{cov}(X,Y). \end{array}}$$

<u>Theorem 4.3</u>: Let $X_1, \dots X_r$ be random variables. We have

$$\operatorname{var}\left(\sum_{k=1}^{r} \alpha_k X_k\right) = \sum_{k=1}^{r} \alpha_k^2 \operatorname{var}(X_k) + \sum_{\substack{k,l=1 \\ k \neq l}}^{r} \alpha_k \alpha_l \operatorname{cov}(X_k, X_l).$$

<u>Proof</u>: Follows directly from Theorem 4.2.

Special case: Let X, Y be discrete and independent.

Then

$$E(X \cdot Y) = \sum_{x,y} x \cdot y \, P(X = x, Y = y)$$

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

$$= \left(\sum_x x P(X = x) \right) \cdot \left(\sum_y y P(Y = y) \right)$$

$$= E(X) \, E(Y).$$

This means $\mathrm{cov}(X, Y) = 0$.

A similar calculation is valid for continuous X, Y.

Remark: For independent X, Y and functions f, g we have

$$E[f(X) g(Y)] = E(f(X)) \, E(g(Y))$$

with the same proof.

As a consequence for <u>independent</u> $X_1, X_2, \dots, X_n$ we have

$$\operatorname{var}(X_1 + \dots + X_n) = \operatorname{var}(X_1) + \dots + \operatorname{var}(X_n)$$

Examples: (i) If $\underline{X} = (X_1, \dots X_r)$ is multinomial we have

$$E(X_k X_\ell) = -np_k p_\ell + n^2 p_k p_\ell$$

and $E(X_k) = np_k$ and $E(X_\ell) = np_\ell$.

We have

$$\mathrm{cov}(X_k, X_\ell) = -np_k p_\ell$$

(ii) If

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}}$$

we have

$$E(X \cdot Y) = \rho \quad \text{and}$$

$$E(X) = E(Y) = 0$$

so

$$\mathrm{cov}(X,Y) = \rho.$$

Method of indicators:

If we can write a random variable X as a sum of indicators we can in many cases compute variances by Theorem 4.3.

If $I \sim \text{Bernoulli}(p)$ then

$$E(I^2) = E(I) = p \quad \text{so}$$

$$\begin{aligned} var(I) &= E(I^2) - E(I)^2 \\ &= p - p^2 \\ &= p(1-p) \end{aligned}$$

If I, J are indicators then

$$\boxed{\begin{aligned} cov(I,J) &= E(I \cdot J) - E(I)E(J) \\ &= P(I=1, J=1) \\ &\quad - P(I=1)P(J=1) \end{aligned}}$$

Example : Let $X \sim \text{HiperGeom}(n, B, N)$.

We wrote

$$X = I_1 + \cdots + I_n .$$

We have

$$\text{var}(X) = \sum_{k=1}^{n} \text{var}(I_k) + \sum_{\substack{k,l=1 \\ k \neq l}}^{n} \text{cov}(I_k, I_l)$$

We know: $I_k \sim \text{Bernoulli}\left(\frac{B}{N}\right)$

so $\text{var}(I_k) = \frac{B}{N} \cdot \left(1 - \frac{B}{N}\right)$.

We have

$$P(I_1 = 1, I_2 = 1) = P(\text{1st \& 2nd ball black})$$

$$= \frac{B}{N} \cdot \frac{B-1}{N-1} .$$

It follows

$$\mathrm{cov}(I_1, I_2) = \frac{B}{N} \cdot \frac{B-1}{N-1} - \left(\frac{B}{N}\right)^2$$

$$= \frac{B}{N}\left[\frac{(B-1)N - B(N-1)}{N(N-1)}\right]$$

$$= \frac{B}{N}\left[\frac{-N+B}{N(N-1)}\right]$$

$$= -\frac{B}{N}\left(1-\frac{B}{N}\right) \cdot \frac{1}{N-1}$$

But by symmetry (I_k, I_l) has the same distribution as (I_1, I_2).

So all covariances are the same.

We have

$$\mathrm{var}(X) = n\frac{B}{N}\left(1-\frac{B}{N}\right) + n(n-1) \times (-1) \times \frac{B}{N}\left(1-\frac{B}{N}\right) \cdot \frac{1}{N-1}$$

$$= n\frac{B}{N}\left(1-\frac{B}{N}\right)\left(1-\frac{n-1}{N-1}\right)$$

$$= n \cdot \frac{B}{N}\left(1-\frac{B}{N}\right)\frac{N-n}{N-1}$$

4.3. Conditional expectation

Idea: If X is a discrete random variable then

$$E[f(X)] = \sum_{x_k} f(x_k) P(X = x_k).$$

The expectation is computed using the distribution. We can use the same idea with the conditional distribution $P(X = x_k \mid B)$ for $P(B) > 0$.

Definition: Let X be a discrete random variable. The conditional expectation of X given B is given by

$$E(X \mid B) = \sum_{x_k} x_k \cdot P(X = x_k \mid B)$$

and

$$E(f(X) \mid B) = \sum_{x_k} f(x_k) P(X = x_k \mid B).$$

Technical note: We understand the existence of $E(X \mid B)$ the same way as for usual expectations.

In most cases B will be of the form $B = \{Y = y_k\}$ for some random variable Y.

<u>Example</u>: Players A and B get 5 cards each from a well shuffled deck of cards. Let X be the number of aces of A and Y the number of aces of B. We have

$Y|_{X=k} \sim \text{HiperGeom}(5, 4-k, 47)$. It follows that

$$E(Y \mid X = k) = 5 \cdot \frac{4-k}{47}$$

We know that for $Z \sim \text{HiperGeom}(n, B, N)$ we have $\text{var}(Z) = n \cdot \frac{B}{N}\left(1 - \frac{B}{N}\right)\frac{N-n}{N-1}$.

so

$$E(Z^2) = \text{var}(Z) + n^2 \cdot \frac{B^2}{N^2}$$

We have

$$E(Y^2 \mid X = k)$$

$$= 5 \cdot \frac{4-k}{47}\left(1 - \frac{4-k}{47}\right) \cdot \frac{47-5}{47-1}$$

$$+ 5^2 \cdot \frac{(4-k)^2}{47^2}$$

There is an alternative way to write the conditional expectation.

If X is discrete and B is an event we compute

$$E[\underbrace{X \cdot \mathbb{1}_B}] = \sum_{x_k} x_k \cdot P(\{X = x_k\} \cap B)$$

$$\uparrow \qquad = (*)$$

This random variable has value x_k with probability $P(\{X = x_k\} \cap B)$ and possibly 0 with probability $P(B \cup \{X = 0\})$

$$(*) = \sum_{x_k} x_k \frac{P(\{X = x_k\} \cap B)}{P(B)} \cdot P(B)$$

$$= P(B) \cdot E(X \mid B)$$

We have

$$E(X \mid B) = \frac{E(X \cdot 1_B)}{P(B)}$$

$$E[f(X) \mid B] = \frac{E[f(X) \cdot 1_B]}{P(B)}.$$

!

<u>Theorem 4.4</u>: Let $\{H_1, H_2, \ldots\}$ be a partition of $(\Omega, \mathcal{F}, P)$. Let X be a discrete random variable with $E(|X|) < \infty$. We have

$$E(X) = \sum_k E(X \mid H_k) \cdot P(H_k)$$

Proof: We compute

$$\sum_k E(X \mid H_k) \cdot P(H_k)$$

$$= \sum_k \left(\sum_{x_\ell} x_\ell \, P(X = x_\ell \mid H_k) \right) P(H_k)$$

$$= \sum_{x_\ell} x_\ell \underbrace{\sum_k P(X = x_\ell \mid H_k) P(H_k)}_{= P(X = x_\ell)}$$

$$= \sum_{x_\ell} x_\ell \, P(X = x_\ell)$$

$$= E(X)$$

We used the law of total probabilities. The statement is the law of total expectations.

Example: We toss a coin until we get r consecutive heads. Tosses are independent and the probability of heads is p. Let X be the number of tosses needed.

Example : if $r = 4$ and we get

$$\underbrace{HTT\ H\ H\ T\ HTT\ HT\ HHHH}_{X = 15}$$

We want $E(X)$. Let

$H_k = \{$the first T appears in position $k\}$.

We have

$$E(X \mid H_k) = r \quad \text{if} \quad k = r+1, \dots$$

and

$$E(X \mid H_k) = k + E(X) \text{ if } k = 1, \dots, r \ ;$$

The law of total expectation gives

$$E(X) = \sum_{k=1}^{\infty} E(X \mid H_k) P(H_k)$$

$$= \sum_{k=1}^{r} (k + E(X)) P(H_k)$$

$$+ \sum_{k=r+1}^{\infty} r \cdot P(H_k)$$

This last expression is a linear equation for $E(X)$. We compute

(i) $$\sum_{k=r+1}^{\infty} r\, P(H_k) =$$

$$= r \cdot \sum_{k=r+1}^{\infty} p^{k-1} \cdot q$$

$$= r \cdot p^r \cdot q \cdot \sum_{k=0}^{\infty} p^k$$

$$= r \cdot p^r \cdot q \cdot \frac{1}{1-p}$$

$$= r \cdot p^r$$

(ii) $$\sum_{k=1}^{r} P(H_k) = q \cdot \sum_{k=1}^{r} p^{k-1}$$

$$= q \cdot \frac{1-p^r}{1-p}$$

$$= 1 - p^r$$

(iii) $$\sum_{k=1}^{r} k \cdot P(H_k) = \sum_{k=1}^{r} k \cdot p^{k-1} \cdot q$$

$$= \frac{d}{dp}\left(p + p^2 + \quad + p^r\right) \cdot q$$

$$= (*)$$

$$(*) = \frac{d}{dp}\left(\frac{p(1-p^r)}{1-p}\right)\cdot q$$

$$= \frac{[(1-p^r) - rpp^{r-1}](1-p) + p(1-p^r)}{(1-p)^2}\cdot q$$

$$= \frac{1-p^r - rp^r + rp^{r+1}}{(1-p)^2}\cdot q$$

$$= \frac{1-p^r - rp^r + rp^{r+1}}{q}$$

Rewrite

$$E(x) = \frac{1-p^r - rp^r + rp^{r+1}}{q}$$

$$+ E(x)\cdot(1-p^r)$$

$$+ r\cdot p^r$$

We have

$$E(x) = r + \frac{1-p^r - rp^r + r\cdot p^{r+1}}{q\cdot p^r}$$

$$= r + \frac{1-p^r + rp^r(p-1)}{q\cdot p^r}$$

$$= \frac{1-p^r}{q\cdot p^r}$$

$$E[f(\underline{X})] = \sum_k E[f(\underline{X}) \mid H_k]\, P(H_k)$$

The proof is identical to the proof we had before.

<u>Definition</u>: We define

$$\mathrm{var}(X \mid B) = E(X^2 \mid B) - [E(X \mid B)]^2$$

and

$$\mathrm{cov}(X, Y \mid B) = E(XY \mid B) - E(X \mid B)\, E(Y \mid B).$$

For continuous random variables we use conditional densities to compute conditional expectations. We have:

<u>Definition</u>: Let X have conditional density $f_{Y \mid X = x}(y)$ given $X = x$.

We define

$$E(Y \mid X = x) = \int_{-\infty}^{\infty} y \, f_{Y \mid X = x}(y) \, dy$$

$$E[f(Y) \mid X = x] = \int_{-\infty}^{\infty} f(y) \, f_{Y \mid X = x}(y) \, dy.$$

Technical note: We define existence the way existence is defined for usual expectations.

Comment: The same definition holds for vectors. We define

$$E[f(\underline{Y}) \mid \underline{X} = \underline{x}]$$

$$= \int_{\mathbb{R}^2} f(y) \, f_{\underline{Y} \mid \underline{X} = x}(y) \, dy.$$

Definition: We define

$$var(Y \mid \underline{X} = \underline{x}) = E(Y^2 \mid \underline{X} = \underline{x}) - [E(Y \mid \underline{X} = \underline{x})]^2$$

and

$\mathrm{cov}(Y_1, Y_2 \mid \underline{X} = \underline{x})$

$$= E(Y_1 \cdot Y_2 \mid \underline{X} = \underline{x})$$

$$- E(Y_1 \mid \underline{X} = x) E(Y_2 \mid \underline{X} = \underline{x})$$

Example: Let

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2 - 2\rho x y + y^2}{2(1-\rho^2)}}$$

We have computed that

$$f_{Y \mid X = x}(y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}}$$

or $Y \mid X=x \sim N(\rho x, 1-\rho^2)$. From this we have

$$E(Y \mid X = x) = \rho x \quad \text{and}$$

$$\mathrm{var}(Y \mid X = x) = 1 - \rho^2.$$