Remark: We already know that $\langle X,X\rangle = E(X^2) \geq 0 \ \forall X$. It can be shown that for any r.v. (not necessarily discrete or cont) X, if $E(X^2) = 0$, then $P(X=0)=1$.

Exc: Show ~~that~~ the above remark when X is either a discrete r.v. or a cont r.v. with a cont pdf.

In other words, show the following:

(a) If X is a discrete r.v. satisfying $E(X^2)=0$, then $P(X=0)=1$.

(b) If X is a cont r.v. with a ~~cot~~ cont pdf, then $E(X^2) > 0$.

Exc: Using the exc stated at the end of Pg 208 and the remark stated at the beginning of Pg 209, show that $\langle \cdot , \cdot \rangle$ (as defined in the last exc of Pg 208) is an inner product on $L^2(\Omega, P)$ making it an inner product space.

Thm (Cauchy-Schwarz Inequality for R.V.s)

Suppose X, Y are jointly distributed r.v.s with finite second moments. Then XY has finite mean and

$$|E(XY)| \leqslant +\sqrt{E(X^2)E(Y^2)}$$

or equivalently, $$(E(XY))^2 \leqslant E(X^2)E(Y^2).$$

Equality holds in either (and hence both) of the above inequalities if and only if $\exists\, \alpha, \beta \in \mathbb{R}$ such that $(\alpha, \beta) \neq (0,0)$ and $P(\alpha X + \beta Y = 0) = 1$.

Proof: Just apply Cauchy-Schwarz inequality on the inner product space $L^2(\Omega, P)$.

Taking $X, Y \in L^2(\Omega, P)$, we observe that $\mathrm{Cov}(X,Y) = \langle X - E(X), Y - E(Y) \rangle$. This observation, together with linearity of expectation, tells us that ~~Cov~~ covariance too is a symmetric $\mathbb{R}$-bilinear form that's clearly not positive-definite.

Exc: Consider the vector subspace

$$L^2_c(\Omega, P) \doteq \{X \in L^2(\Omega, P) : E(X) = 0\}$$

of $L^2(\Omega, P)$ consisting of all centred (i.e. mean = 0) r.v.s (with finite 2nd moments).

(i) Show that $E : L^2(\Omega, P) \longrightarrow \mathbb{R}$ defined by

$$X \longmapsto E(X)$$

is a well-defined, surjective $\mathbb{R}$-linear functional.

(ii) Show that $L^2_c(\Omega, P)$ is the kernel of the linear functional described in (i).

(iii) Show that $(X, Y) \longmapsto \mathrm{Cov}(X, Y)$ is an inner product on $L^2_c(\Omega, P)$.

Exc: For $X, Y \in L^2(\Omega, P)$, show the following:

(i) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$ if and only if X and Y are uncorrelated.

(ii) $\mathrm{Cov}(X, Y) = \dfrac{\mathrm{Var}(X+Y) - \mathrm{Var}(X-Y)}{4}$.

(iii) X and Y are uncorrelated if and only if $\mathrm{Var}(X+Y) = \mathrm{Var}(X-Y)$.

Remark: As stressed before, $L^2(\Omega, P)$ is the collection of all r.v.s defined on Ω and ~~and~~ having finite 2^{nd} moments such that the equality of two r.v.s is understood in the almost sure sense, i.e.,

$$X = Y \Leftrightarrow P(X=Y)=1.$$

In a similar manner, we can define the space $L^1(\Omega, P)$ as the collection of all r.v.s defined on Ω and having finite mean such that the equality of two r.v.s is defined in the almost sure sense. In fact, for any $p \in \mathbb{N}$, the space $L^p(\Omega, P)$ can be defined analogously.

Exc: Suppose $p, q \in \mathbb{N}$ with $p \leq q$. Show that

$$L^q(\Omega, P) \subseteq L^p(\Omega, P).$$

In particular

$$L^2(\Omega, P) \subseteq L^1(\Omega, P).$$

Remark: L^p spaces are also called Lebesgue spaces.

We know that equality holds in the Cauchy-Schwarz inequality (see Pg 210) if and only if $\exists\ (\alpha, \beta) \in \mathbb{R}^2 \setminus \{(0,0)\}$ such that $P(\alpha X + \beta Y = 0) = 1$.

<u>Exc:</u> Show that equality holds in the Cauchy-Schwarz inequality if and only if either $X \equiv 0$ (i.e., $P(X=0)=1$) or $\exists\ \gamma \in \mathbb{R}$ such that $P(Y = \gamma X) = 1$.

<u>Questions:</u> (i) When will $E(XY) = +\sqrt{E(X^2)E(Y^2)}$ hold?

(ii) When will $E(XY) = -\sqrt{E(X^2)E(Y^2)}$ hold?

In order to answer the above questions, we have to go back to the proof of Cauchy-Schwarz inequality. This proof (in this special case of r.v.s) goes as follows: the quadratic expression $q(t)$

$$= E[(Y - tX)^2] = E(X^2)t^2 - 2E(XY)t + E(Y^2) \geq 0$$

for all $t \in \mathbb{R}$ and hence the discriminant should be nonpositive, from which the Cauchy-Schwarz inequality follows.

Answer to Question (i) in Pg 213

Suppose $\quad E(XY) = +\sqrt{E(X^2)\,E(Y^2)}$.

Then
$$q(t) = E\left[(Y - tX)^2\right]$$
$$= E(X^2)\,t^2 - 2\,E(XY)\,t + E(Y^2)$$
$$= E(X^2)\,t^2 - 2\,\overset{+}{\sqrt{E(X^2)E(Y^2)}}\;t + E(Y^2)$$
$$= \left(+\sqrt{E(X^2)}\;t - \overset{+}{\sqrt{E(Y^2)}}\right)^2.$$

Suppose $P(X=0) < 1$. Then $E(X^2) > 0$.

In this case, $\quad q\left(\dfrac{+\sqrt{E(Y^2)}}{+\sqrt{E(X^2)}}\right) = 0$, i.e.,

$$q(\gamma) = E\left[(Y - \gamma X)^2\right] = 0, \quad \text{where}$$

$$\gamma = \frac{\overset{+}{\sqrt{E(Y^2)}}}{\overset{+}{\sqrt{E(X^2)}}} \in [0, \infty)$$ under the hypothesis of Cauchy-Schwarz inequality.

We have shown: $E(XY) = +\sqrt{E(X^2)E(Y^2)}$ and $P(X=0)<1$

$$\Rightarrow q(\gamma) = E\left[(Y-\gamma X)^2\right] = 0 \quad \text{for some } \gamma \in [0,\infty)$$

$$\Rightarrow P[Y = \gamma X] = 1 \quad \text{for some } \gamma \in [0,\infty)$$

In other words,

$$E(XY) = +\sqrt{E(X^2)E(Y^2)} \Rightarrow P(X=0)=1 \text{ or } P[Y=\gamma X] = 1 \text{ for some } \gamma \in [0,\infty).$$

Conversely, if $P(X=0)=1$, then

$$E(XY) = 0 = +\sqrt{E(X^2)E(Y^2)}.$$

On the other hand, if $\exists\, \gamma \in [0,\infty)$ such that $P[Y=\gamma X]=1$, then

$$E(XY) = E(X\gamma X) = \gamma E(X^2)$$

and

$$+\sqrt{E(X^2)E(Y^2)} = +\sqrt{E(X^2)E(\gamma^2 X^2)} = +\sqrt{\gamma^2 (E(X^2))^2}$$

$$= \gamma E(X^2) \quad [\because \gamma \in [0,\infty)]$$

yielding $E(XY) = +\sqrt{E(X^2)E(Y^2)}$.

Conclusion: Under the hypothesis of Cauchy-Schwarz inequality (see Pg 210), $E(XY) = +\sqrt{E(X^2)\,E(Y^2)}$ holds if and only if either $P(X=0)=1$ or $\exists\ \gamma \in [0,\infty)$ such that $P[Y=\gamma X]=1$.

The above conclusion answers Question (i) of Pg 213.

Exc: Answer Question (ii) of Pg 213 by showing the following: under the hypothesis of Cauchy-Schwarz inequality, $E(XY) = -\sqrt{E(X^2)\,E(Y^2)}$ holds if and only if either $P(X=0)=1$ or $\exists\ \gamma \in (-\infty, 0]$ such that $P[Y=\gamma X]=1$.

Remarks: ① One can solve the above exercise either by following the proof of the conclusion above or with the help of the exercise given in Pg 213.

② Using symmetry, one can switch the roles of X and Y in the above exc, or the conclusion above or the exc given in Pg 213.

Def^n: Suppose X, Y are jointly distributed nondegenerate r.v.s having finite second moments. Then the correlation or the correlation coefficient of X and Y is defined as

$$\mathrm{Corr}(X,Y) = \rho(X,Y) = \rho_{X,Y} := \frac{\mathrm{Cov}(X,Y)}{+\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}.$$

Remark: The conditions given in the above def^n ensure that $\rho_{X,Y}$ is well-defined. By the remark stated at the end of Pg 175, $\mathrm{Var}(X) \neq 0$, $\mathrm{Var}(Y) \neq 0$.

Properties of Correlation Coefficient

(0) Suppose X is a nondegenerate r.v. with finite second moment, then $\rho(X,X) = 1$.

Proof: $$\rho(X,X) = \frac{\mathrm{Cov}(X,X)}{+\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(X)}} = \frac{\mathrm{Var}(X)}{\mathrm{Var}(X)} = 1$$

(again, since X nondegenerate $\Rightarrow \mathrm{Var}(X) > 0$).

(1) (Symmetry) Suppose X, Y are jointly distributed nondegenerate r.v.s having finite 2^{nd} moments. Then
$$\mathrm{Corr}(X,Y) = \mathrm{Corr}(Y,X).$$