Example 0.1. *suppose x has a symmetric dist about zero. ie. $P(x < -t) = P(x > t)$ for all $t \geqslant 0$. suppose X and Z are independent. let $Y = x^2 + Z$, find P_x.*

$$
\begin{aligned}
&x = x \cdot I[x > 0] + X \cdot I[x = 0] + X \cdot I\,[x < 0] \\
&\operatorname{cov}(x, y) = \operatorname{cov}\left(x, x^2 + Z\right) = \operatorname{cov}\left(x \cdot x^2\right) + \operatorname{cov}(x, z) \\
&\operatorname{cov}\left(x, x^2\right) = \operatorname{cov}\left(x_+ + x_-, x^2\right) = \operatorname{cov}\left(x_+, x^2\right) + \operatorname{cov}\left(X_-, X^2\right)
\end{aligned}
$$

By symmetric, $x_- = -x_+$, Thus $\left(x_-, x^2\right) = \left(-x_+, x^2\right)$

$$
\begin{aligned}
&\operatorname{cov}\left(x, x_-^2\right) = \operatorname{cov}\left(-x_+, x^2\right) = -\operatorname{cov}\left(x_+, x^2\right) \\
&\operatorname{cov}\left(x_+, x^2\right) - \operatorname{cov}\left(x_+, x^2\right) = 0.
\end{aligned}
$$

1 Bivariate Normal distribution.

R.V.S X and Y has a bivariate normal dist. if their joint density is given by

$$
f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-p^2}} \exp\left(-\frac{1}{2(1-p^2)}(\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2p\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2)\right)
$$

We will Summarize this as

$$
\left[\begin{array}{c} x \\ y \end{array}\right] \sim N\left(\left[\begin{array}{c} \mu_x \\ \mu_y \end{array}\right], \left[\begin{array}{cc} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{array}\right]\right.
$$

property

(1) marginal and Conditional Dist.

$$
\begin{aligned}
&x \sim N\left(\mu_x, \sigma_x^2\right) \\
&Y \mid X \sim N(\underbrace{\mu_y + \rho\frac{\sigma_y}{b_x}\left(x - \mu_x\right)}_{m(x)=E[Y|X]}, b_y^2\left(1 - p^2\right))
\end{aligned}
$$

(2) $\operatorname{Corr}(X, Y) = \rho$

Proof. $\operatorname{cov}(X, Y) = \sigma_{xy} = E\left(\left(x - \mu_x\right)\left(y - \mu_y\right)\right)$

$$
\begin{aligned}
&= E_x\left(E\left(\left(x - \mu_x\right)\left(y - \mu_y\right) \mid x\right)\right) \\
&= E_x\left(\left(x - \mu_x\right) E\left(y - \mu_y \mid x\right)\right) \\
&= E_x\left(\left(x - \mu_x\right) \cdot \rho\frac{\sigma_y}{\sigma_x}\left(x - \mu_x\right)\right) \\
&= E_x\left(\rho\frac{\sigma_y}{\sigma_x}\left(x - \mu_x\right)^2\right) \\
&= \rho\frac{\sigma_y}{\sigma_x} \cdot \sigma_x^2 = \rho\sigma_y \cdot \sigma_x
\end{aligned}
$$

□

(3) regression function.
Define $\alpha=\mu_y, \beta=\rho \frac{\sigma_y}{\sigma_x}, \sigma_\epsilon=\sigma_y \sqrt{1-p^2}$ and $\gamma=\mu_y-\beta \mu_x$
LBE of Y given X is the conditional mean.

$$
\begin{aligned}
& m(x)=E(Y \mid X)=\mu_Y+\left(\rho \frac{\sigma_y}{\sigma_x}\right)\left(X-\mu_x\right)=\gamma+\beta x \\
& V^2(x)=V(Y \mid X=x)=\sigma_y^2\left(1-\rho^2\right)=\sigma_\epsilon^2 \leftarrow \text { constant }
\end{aligned}
$$

(4) (x, y) are independent if and only if $\rho=0$.
proof $f(x, y)=f(x) \cdot f(y)$.

Definition 1.1. *The standard deviation line is dentine as the following*

$$
y=\mu_y+\frac{\sigma_y}{\sigma_x}\left(x-\mu_x\right) \text { or } \frac{y-\mu_y}{\sigma_y}=\frac{x-\mu_x}{\sigma_x}
$$

(5) Let $u=a x+b y+c$ and $V=d X+e^y+f$, Then (u, v) is bivariate normal whose dist is completely specified by $\mu_u, \mu_v, \sigma_u^2, \sigma_v^2, \sigma_{u v}$
Note: $(X, Y) \sim$ Bivariate normal $\Rightarrow X \sim$ Normal, Y $\sim$ Normal

Theorem 1.2. *Cauchy swartzs inequality*

$$
|E(X Y)| \leq E(|X Y|) \leq E\left(X^2\right)^{\frac{1}{2}} E\left(Y^2\right)^{1 / 2}
$$

Corollary: $\left|\rho_{x y}\right| \leq 1$

Proof. let $a=x-\mu_x$ and $b=y-\mu_y$

$$
\begin{aligned}
& |\operatorname{cov}(a, b)|=|E(a, b)| \leq E\left(a^2\right)^{1 / 2} E\left(b^2\right)^{1 / 2}=\left(\sigma_x^2\right)^{1 / 2} \cdot\left(\sigma_y^2 y\right)^{1 / 2} \\
& \Rightarrow \frac{|\operatorname{cov}(x, y)|}{\sigma_x b_y} \leq 1
\end{aligned}
$$

□

Theorem 1.3. *Jensen's inequality. For any riv. x, if $g(x)$ is a convex function. then $E(g(x)) \geqslant g(E(x))$ if and only if $g(x)=a+b x$.*

Proof. let $((x)$ be the tangent line tog (x) at $E(x)$.

$$
E(g(x)) \geqslant E(l(x))=((E(x))=g(E(x))
$$

□

Example 1.4. *Let g in $)=x^2$, then $E\left(x^2\right)$ y $E(x)^2$*