Stat 134 lec 41

warmup 11:00–11:10

let X, Y std bivariate normal, $\rho > 0$

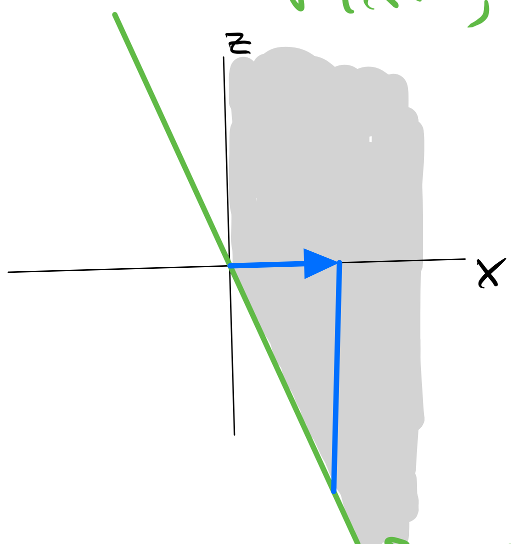
Find $P(X>0, Y>0)$

Hint

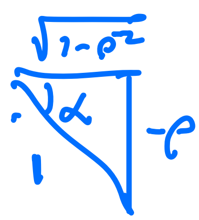
$$P(X>0, Y>0) = P(X>0, \rho X + \sqrt{1-\rho^2}\, Z > 0)$$

$$= P\left(X>0, Z > \frac{-\rho}{\sqrt{1-\rho^2}} X\right)$$

note X and Z are uncorrelated so joint (X, Z) is a symmetric bell over x, z plane,



$$z = \frac{-\rho}{\sqrt{1-\rho^2}} x$$



$$\tan\alpha = \frac{-\rho}{\sqrt{1-\rho^2}} \Rightarrow \alpha = \tan^{-1}\left(\frac{-\rho}{\sqrt{1-\rho^2}}\right)$$

$$P\left(X>0, Z > \frac{-\rho}{\sqrt{1-\rho^2}} X\right) = \boxed{\frac{90 + |\alpha|}{360}}$$

last time

RRR week schedule

M Review (post questions on b-Courses),

W OFF

F Review (post questions on b-courses)

OH after class on zoom

Sec 6.5. Bivariate Normal

Def (Standard Bivariate Normal Distribution)

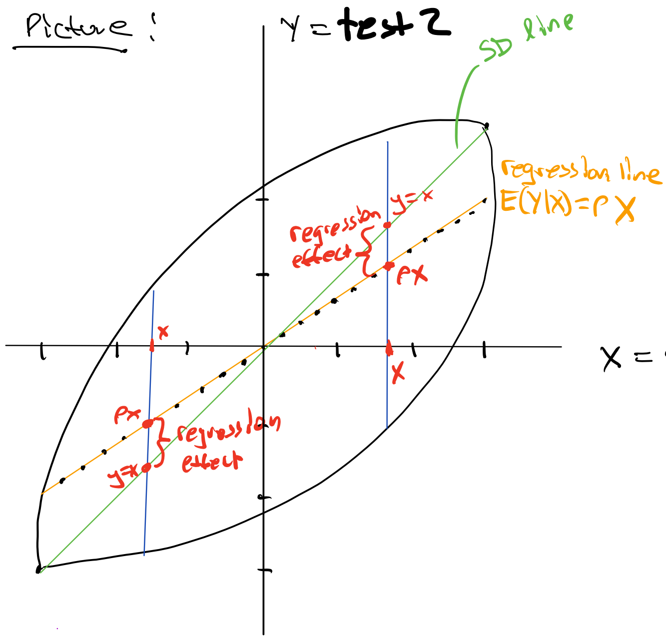
let X, Z iid $N(0,1)$, $-1 \le \rho \le 1$

$$Y = \rho X + \sqrt{1-\rho^2}\, Z \sim N(0,1)$$

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

regression **effect**

Picture:



Regression effect,

Corr(test 1, test 2) = .6

If 1 SD above mean on test 1 then on average you will be less than 1 SD above average on test 2.

(regression line is less steep than SD line).

Today ① MGF of bivariate normal

② sec 6.5 properties of bivariate normal

③ sec 6.4 Total Variance Decomposition

① MGF of bivariate normal

The single and multivariate MGF is defined as:

$$M_Y(t) = E\left(e^{tY}\right) \quad \text{single variable MGF}$$

$$M_{(X,Y)}(s,t) = E\left(e^{sX+tY}\right) \quad \text{multivariate MGF}$$

Show that $M_{(X,Y)}(s,t) = M_X(s)\,M_Y(t)$ iff X, Y are independent.

$$M_{(X,Y)}(s,t) = E\left(e^{sX+tY}\right)$$

$$= E\left(e^{sX} \cdot e^{tY}\right)$$

iff X, Y indep.

$$= E\left(e^{sX}\right) E\left(e^{tY}\right)$$

$$= M_X(s) \cdot M_Y(t)$$

$E(AB) = E(A)E(B)$ iff A, B are indep.

Here
$A = e^{sX}$
$B = e^{tY}$

Thm Let (X,Y) be standard bivariate normal.
The MGF of (X,Y) is

$$M_{(X,Y)}(s,t) = e^{\frac{s^2}{2} + \frac{t^2}{2} + st\rho}$$

Pf/ $M_{X,Y}(s,t) = E\left[e^{sX+tY}\right]$

$$= E\left[e^{sX + t(\rho X + \sqrt{1-\rho^2}Z)}\right]$$

$$= E\left[e^{(s+t\rho)X} \cdot e^{t\sqrt{1-\rho^2}Z}\right]$$

$$\overset{\text{independence}}{=} E\left[e^{(s+t\rho)X}\right] E\left[e^{t\sqrt{1-\rho^2}Z}\right]$$

$$= M_X(s+t\rho) \cdot M_Z(t\sqrt{1-\rho^2})$$

Recall that $X \sim N(0,1)$ so $M_X(a) = e^{\frac{a^2}{2}}$

Finish proof.

$$= e^{\frac{(s+t\rho)^2}{2}} \cdot e^{\frac{(t\sqrt{1-\rho^2})^2}{2}}$$

$$= e^{s^2/2 + st\rho + \cancel{\frac{t^2\rho^2}{2}}} \cdot e^{\frac{t^2}{2} - \cancel{\frac{t^2\rho^2}{2}}}$$

$$= \boxed{e^{s^2/2 + t^2/2 + st\rho}}$$

② Properties of Bivariate Normal

Recall that if 2 RVs X, Y are independant, $Cov(X,Y)=0$

$\Rightarrow Corr(X,Y)=0$

However the converse isn't true in generall. ($Corr(X,Y)=0 \not\Rightarrow X,Y$ indep.)

Let $X \sim N(0,1)$

$Y = X^2$

X and Y are dependent

we show X, Y are uncorrelated.

$Cov(X,Y) = \underset{\overset{\shortparallel}{E(X^3)}}{E(XY)} - E(X)\underset{\overset{\shortparallel}{0}}{E(Y)}$

Real world example

X = height

Y = $(\text{height})^2$

If you make a scatter plot of height vs $(\text{height})^2$ it will look uncorrelated even though height and $(\text{height})^2$ are dependent

Find $E(X^3)$

$$= \int_{-\infty}^{\infty} x^3 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx = 0$$

since $\frac{x^3}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ is an odd function

or

$M_X(t) = e^{t^2/2}$

$$M_X'''(0) = t e^{t^2/2} + 2t e^{t^2/2} + t^3 e^{t^2/2}\Big|_{t=0} = 0$$

$\Rightarrow E(X^3)=0$

$\Rightarrow Corr(X,Y)=0$

<u>Thm</u> If (X,Y) is std bivariate normal then $\rho = \mathrm{Corr}(X,Y) = 0$ iff X, Y are independent.

Pf/ From the warm up we know for any joint distribution (X,Y),

$$M_{(X,Y)}(s,t) = M_X(s)\, M_Y(t) \quad \text{iff } X, Y \text{ are independent.}$$

Since (X,Y) is std bivariate normal,

$$M_{(X,Y)}(s,t) = e^{s^2/2 + t^2/2 + st\rho}$$

X,Y indep $\quad M_{(X,Y)}(s,t) = M_X(s)\, M_Y(t) \quad \Rightarrow e^{st\rho} = 1$

$$e^{s^2/2 + t^2/2 + st\rho} = e^{s^2/2} \cdot e^{t^2/2} \quad \Rightarrow \boxed{\rho = 0}$$

Conversely, if $\rho = 0$

$$\underbrace{e^{s^2/2 + t^2/2 + st\rho}}_{M_{(X,Y)}(s,t)} = e^{s^2/2 + t^2/2} = \underbrace{e^{s^2/2}}_{M_X(s)} \cdot \underbrace{e^{t^2/2}}_{M_Y(t)} \Rightarrow X, Y \text{ indep.}$$

Recall from lec 30 that the sum of <u>independent</u> normal random variables is normal.

What about <u>dependent</u> normal random variables?

recall,

$M_Y(t) = E(e^{tY}) = M_{tY}(1)$ single variable MGF

$M_{(X,Y)}(s,t) = E(e^{sX+tY}) = M_{sX+tY}(1)$ multivariable MGF

recall,

$Z \sim N(\mu,\sigma^2)$ iff $M_Z(a) = e^{\mu a} e^{\sigma^2 \frac{a^2}{2}}$

Thm Let $X, Y \sim N(0,1)$ and $\mathrm{corr}(X,Y) = \rho$.
(X,Y) is std bivariate normal iff
$sX + tY$ is normal for all constants s,t.

Pf/
$\Rightarrow$: Suppose X, Y std bivariate normal.
(i.e. $Y = \rho X + \sqrt{1-\rho^2}\, Z$ for X, Z iid $N(0,1)$)

$$sX + tY = sX + t(\rho X + \sqrt{1-\rho^2}\, Z)$$
$$= (s + t\rho)X + t\sqrt{1-\rho^2}\, Z$$

is normal since X, Z are indep normals

$\Leftarrow$: Suppose $sX+tY$ is normal,

note $E(sX+tY) = sE(X) + tE(Y) = 0$

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

$Var(sX+tY) = Var(sX) + Var(tY) + 2\,Cov(sX, tY)$

$= s^2 + t^2 + 2st\rho$

$2\,Cov(sX,tY) = 2st\,Cov(X,Y)$, $Cov(X,Y) = \rho$

Hence, $sX+tY \sim N(0, s^2+t^2+2st\rho)$

Recall,

For $Z \sim N(\mu, \sigma^2) \Rightarrow M_Z(a) = e^{\mu a} e^{\sigma^2 \frac{a^2}{2}}$

Then $M_{(X,Y)}(s,t) = M_{sX+tY}(1) = e^{(s^2+t^2+2st\rho)\frac{1^2}{2}}$

normal

σ^2

$= e^{\frac{s^2}{2} + \frac{t^2}{2} + st\rho}$ $\leftarrow$ MGF of (X,Y)

$\Rightarrow (X,Y)$ is standard bivariate normal $\square$

We dont need to restrict ourselves to $X, Y \sim N(0,1)$

Corollary — Proof at end of lecture

Let $U \sim N(\mu_U, \sigma_U^2)$
$V \sim N(\mu_V, \sigma_V^2)$
and $Corr(U,V) = \rho$
$(U,V) \sim BV(\mu_U, \mu_V, \sigma_U^2, \sigma_V^2, \rho)$ iff
$sU + tV$ is normal for all constants s,t.

ex

Let M be a students score on the midterm of a class and F the student's score on the final of the same class.

Suppose $(M, F) \sim BV(70, 65, 8^2, 10^2, 0.6)$.

$\mu_M \quad \mu_F \quad \sigma_M^2 \quad \sigma_F^2 \quad \rho$

Find the chance that the student scores higher on the final than the midterm.

Hint $P(F > M) = P(F - M > 0)$

what distribution is $F - M$? — Normal! (Since (M, F) is Bivariate Normal)

$E(F-M) = 65 - 70 = -5$

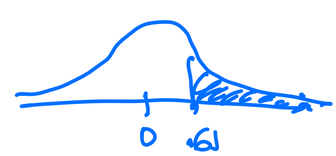
$$Var(F-M) = Var(F) + Var(M) - 2\underbrace{Cov(F,M)}_{= Corr(F,M)\,SD(M)\,SD(F)}$$

$$= 10^2 + 8^2 - 2(.6)(8)(10) = 68$$

$SD(F-M) = \sqrt{68}$

$\Rightarrow F - M \sim N(-5, 68)$

$$\Rightarrow \frac{0 - E(F-M)}{SD(F-M)} = \frac{0 - (-5)}{\sqrt{68}} = \boxed{.61}$$



$P(F > M) = P(F - M > 0) = 1 - \Phi(.61) = .27$

Appendix

Corollary

Let $U \sim N(\mu_U, \sigma_U^2)$
$V \sim N(\mu_V, \sigma_V^2)$
and $\text{Corr}(U,V) = \rho$
$(U,V) \sim BV(\mu_U, \mu_V, \sigma_U^2, \sigma_V^2, \rho)$ iff
$sU + tV$ is normal for all constants s,t.

Pf/ let $X = \dfrac{U-\mu_U}{\sigma_U}$ $\quad (U = \mu_U + X\sigma_U)$

$Y = \dfrac{V-\mu_V}{\sigma_V}$ $\quad (V = \mu_V + Y\sigma_V)$

$(U,V) \sim BV(\mu_U, \mu_V, \sigma_U^2, \sigma_V^2, \rho)$
iff $(X,Y) \sim BV(0, 0, 1, 1, \rho)$

$$sU + tY = s(\mu_U + X\sigma_U) + t(\mu_V + Y\sigma_V)$$
$$= \underbrace{s\sigma_U}_{=a} X + \underbrace{t\sigma_V}_{=b} Y + \underbrace{sx\sigma_U + ty\sigma_V}_{\text{constant}}$$

this is normal for all constants s,t
iff $aX + bY$ is normal for all constants a,b.

$\square$