STAT 24400
Lecture 5
January 19, 2016

We've talked about expectations, and we've talked about joint distributions. What are expectations of joint distributions $E[h(X,Y)]$?

As with a single variable:

we could ① Let $Z = h(X,Y)$, a random var, and then find $f_Z(z)$, find $\int_{-\infty}^{\infty} z f_Z(z)\,dz$

(this is hard)

or

② Find $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\,f(x,y)\,dx\,dy$

In general, we sidestep the issue and define

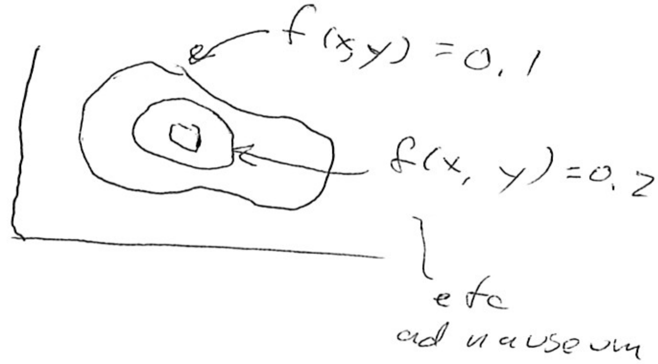
$$E(h(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\,f(x,y)\,dx\,dy$$

$$E(h(X_1, X_2, \dots, X_n)) = \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n\text{-fold}} h(x_1,\dots,x_n)\,f(x_1,\dots x_n)\,dx_1\dots dx_n$$

Describing Multivariate Distributions

One Way: $f(x, y)$

Hard to graph, and for more than 2D **very hard** to graph.



Another Way: single number summaries, covariance, correlation

---

Recall $E(h(X, Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f(x,y)\, dx\, dy$

for $h(x, y) = x + y$

$$E(X+Y) = \iint (x+y) f(x,y)\, dx\, dy$$

$$= \iint x f(x,y)\, dx\, dy + \iint y f(x,y)\, dx\, dy$$

$$= \int x \underbrace{\left[\int f(x,y)\, dy\right]}_{f_X(x)} dx + \int y \underbrace{\left[\int f(x,y)\, dx\right]}_{f_Y(y)} dy$$

$$= \int x f_X(x)\, dx + \int y f_Y(y)\, dy$$

$$= E(X) + E(Y)$$

More Generally: For any random variables $X_1, X_2, \ldots, X_n$

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

hence: $$E\left(\sum_{i=1}^{n} h_i(X_i)\right) = \sum_{i=1}^{n} E(h_i(X_i))$$

In particular, expectations of linear functions are linear functions of <u>marginal</u> <u>expectations</u>:

$$E(aX + bY) = aE(X) + bE(Y)$$

marginal, univariate

Since marginal distributions do not in general determine a bivariate distribution, we <u>cannot</u> describe bivariate dists (only) with <u>expectations</u> of linear funcs

need more:

Covariance of X and Y

$$\text{cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

$$[\,] = XY - \mu_x Y - \mu_y X + \mu_x \mu_y$$

$$E[\,] = E[XY] - \mu_x \underbrace{E(Y)}_{\mu_y} - \mu_y \underbrace{E(X)}_{\mu_x} + \mu_x \mu_y$$

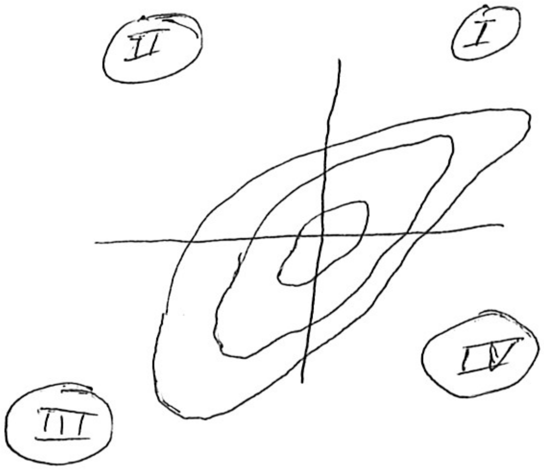
$$= E[XY] - \mu_x \mu_y$$

$$\boxed{\text{cov}(X, Y) = E[XY] - \mu_x \mu_y}$$

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f(x,y)\, dx\, dy$$

note that if $E(X) = E(Y) = 0$

$$\text{Cov}(X,Y) = E(XY)$$



$E(XY) > 0$ if more prob. in (I + III) than (II + IV)

note further that $\text{Cov}(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$ implies that

$$\text{Cov}(X,X) = \text{Var}(X)$$

---

other properties:

$$\text{Cov}(a + bX, c + dY) = bd\,\text{Cov}(X,Y)$$

If X, Y indep. $\text{Cov}(X,Y) = 0$

But

Can have $\text{Cov}(X,Y) = 0$ with X, Y dependent

Correlation: "scale free" covariance

$$\text{corr}(X,Y) = \rho_{XY} = \text{Cov}\left(\frac{X-\mu_x}{\sigma_x}, \frac{Y-\mu_y}{\sigma_Y}\right) = \frac{\text{Cov}(X,Y)}{\sigma_x \sigma_y}$$

also called "Pearson's Correlation Coefficient"

Example:

$$f(x,y) = \begin{cases} \frac{4}{5}(xy+1) & \begin{matrix} 0<x<1 \\ 0<y<1 \end{matrix} \\ 0 & \text{otherwise} \end{cases}$$

Is it really a distribution? Let's check!

$$\int_0^1\int_0^1 f(x,y)\,dx\,dy = \int_0^1 \underbrace{\left(\int_0^1 \frac{4}{5}(xy+1)\,dx\right)}_{f_Y(y)} dy$$

$$= \int_0^1 \underbrace{\frac{4}{5}\left(\frac{1}{2}y+1\right)}_{f_Y(y)} dy = \frac{4}{5}\left(\frac{y^2}{4}+y\right)\Big|_0^1 = 1$$

$$E(Y) = \int_0^1 y\, f_Y(y)\,dy = \frac{4}{5}\int_0^1 \left(\frac{1}{2}y^2+y\right)dy = \frac{8}{15}$$

similarly $E(X) = 8/15$

$$E(XY) = \int_0^1\int_0^1 xy\, f(x,y)\,dx\,dy = \int_0^1\left[\int_0^1 \frac{4}{5}(x^2y^2+xy)\,dx\right]dy$$

$$= \int_0^1 \frac{4}{5}\left[\frac{y^2}{3}+\frac{y}{2}\right]dy = \frac{4}{5}\left[\frac{y^3}{9}+\frac{y^2}{4}\right]_0^1 = \frac{13}{45}$$

$$\Rightarrow \mathrm{Cov}(X,Y) = \frac{13}{45} - \left(\frac{8}{15}\right)\left(\frac{8}{15}\right) = \frac{1}{225}$$

Can find ρ_{xy}. But what does the magnitude of covariance mean?

To interpret magnitude of Cov, consider $X+Y$.

$$E(X+Y) = E(X) + E(Y) = \mu_X + \mu_Y$$

$$\begin{aligned}
Var(X+Y) &= E\left[(X+Y) - E(X+Y)\right]^2 \quad \text{(def)} \\
&= E\left[(X+Y) - (\mu_X + \mu_Y)\right]^2 \\
&= E\left[(X-\mu_X) + (Y-\mu_Y)\right]^2 \\
&= E\left[(X-\mu_X)^2 + (Y-\mu_Y)^2 + 2(X-\mu_X)(Y-\mu_Y)\right] \\
&= Var(X) + Var(Y) + 2\,Cov(X,Y)
\end{aligned}$$

General Interp. of Covariance:

It is a <u>correction factor</u> for finding variances of sums

$$\Rightarrow \boxed{Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)}$$

So:

$$Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$$

if $Cov(X,Y) = 0$ ("X, Y uncorrelated")

then $Var(X+Y) = Var(X) + Var(Y)$

More generally:

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + \sum_{i}\sum_{j}_{i \neq j} Cov(X_i, X_j)$$

using

$$Var(aX) = a^2 Var(X)$$
$$Cov(aX, bY) = ab\,Cov(X,Y)$$

we get

$$Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum\sum_{i \neq j} a_i a_j Cov(X_i, X_j)$$

If each pair X_i and X_j $i \neq j$ are uncorrelated

$$Var\left(\sum a_i X_i\right) = \sum a_i^2 Var(X_i)$$

Let's continue with the case

$Cor(X_i, X_j) = 0$ iff $i \neq j$, so

$$Var(\Sigma a_i X_i) = \Sigma a_i^2 Var(X_i)$$

introduce a new random var

$$\bar{X} = \Sigma \frac{1}{n} X_i \qquad (\text{so } a_i = \frac{1}{n})$$

Then $Var(\bar{X}) = \frac{1}{n^2} \Sigma Var(X_i)$

now suppose the X_i are identically distributed, so

$$Var(X_i) = \sigma^2 \text{ for all } i$$

then

$$\underline{\underline{Var(\bar{X}) = \frac{\sigma^2}{n}}}$$

Note that $E(\bar{X}) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \mu_x$

Recast more carefully in terms of limits, this is the Law of Large Numbers

(see Rice, p 178)

We will return to $\bar{X}$ after a short excursion

Summary

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

$$E(h(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f(x,y)\,dx\,dy$$

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

($= 0$ if X, Y independent)

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

$$Var(X_1 + \cdots + X_n) = \sum_{i=1}^{n} Var(X_i) + \sum_{i \neq j}\sum Cov(X_i, X_j)$$

Indep. Case

$$Var(X+Y) = Var(X) + Var(Y)$$

$$Var(\sum X_i) = \sum_{i=1}^{n} Var(X_i)$$

$$Var(\bar{X}) = \frac{\sigma^2}{n} \quad \text{if } Var(X_1) = Var(X_2) = \cdots = \sigma^2$$

$$E(\bar{X}) = \mu \quad \text{if } E(X_1) = E(X_2) = \cdots = \mu$$

Now let's digress, for a moment...

<u>Def</u> The r^{th} <u>moment</u> of random variable X is $E(X^r)$. (if $E(X^r)$ exists.)

We have already looked at the first moment, $E(X)$, and the second moment $E(X^2)$.

<u>Def</u> The r^{th} <u>central</u> <u>moment</u> of rand. var X is

$$E\left[(X - E(X))^r\right]$$

1^{st} central moment: zero (the mean)
2^{nd} central moment: variance, etc.

It turns out that there is a great trick for dealing with moments – the <u>moment-generating function</u> (mgf)

$$M(t) = E\left[e^{tX}\right]$$

Why care about the moment-generating function? (we will consider the cont. case)

$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$M'(t) = \frac{d}{dt}\int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \int_{-\infty}^{\infty} x e^{tx} f(x)\,dx$$

For $t = 0$

$$M'(0) = \int_{-\infty}^{\infty} x f(x)\,dx = E(X)$$

$$M''(t) = \frac{d}{dt}\int_{-\infty}^{\infty} x e^{tx} f(x)\,dx$$

$$= \int_{-\infty}^{\infty} x^2 e^{tx} f(x)\,dx$$

$$M''(0) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = E(X^2)$$

without proof, it turns out that if $M(t)$ exists in an open interval containing zero $M^{(r)}(0) = E(X^r)$

Moreover

If the moment-generating function exists for t in an open interval containing zero, it uniquely determines the probability distribution.

So we can work with mgf's if we want to, instead of pdf's or cdf's

$$\text{mgf} \rightleftharpoons \text{pdf} \rightleftharpoons \text{cdf}$$

The properties of expectations that we already know enable us to deduce important properties of the mgf.

say X has mgf $M_X(t)$ and $Y = a + bX$

$$\begin{aligned} M_Y(t) &= E(e^{tY}) \\ &= E(e^{at+btX}) \\ &= E(e^{at}e^{btX}) \\ &= e^{at}E(e^{btx}) \end{aligned}$$

$$\boxed{M_Y(t) = e^{at}M_X(bt)}$$

Say X and Y are indep. rand. variables with mgf's M_X and M_Y.
Let $Z = X + Y$

$$M_Z(t) = E(e^{tZ})$$
$$= E(e^{tX + tZ})$$
$$= E(e^{tX} e^{tY})$$

because X and Y are indep

$$= E(e^{tX}) E(e^{tY})$$

$$\boxed{M_Z(t) = M_X(t) M_Y(t)}$$

Example: The Standard Normal Dist.

$$M(t) = E(e^{tX}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2} + tx} dx$$

note that

$$\frac{x^2}{2} - tx = \frac{1}{2}(x^2 - 2tx + t^2) - \frac{t^2}{2}$$
$$= \frac{1}{2}(x-t)^2 - t^2/2$$

so $X \sim \mathcal{N}(0,1)$, as we were saying

$$M(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2} + tx} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-t)^2 - t^2/2} dx$$

$$e^{t^2/2} \underbrace{\left(\frac{1}{\sqrt{2\pi}} \right) \int_{-\infty}^{\infty} e^{-(x-t)^2/2} dx}_{= 1}$$

(set $u = x - t$ change of vars)

$$\boxed{M(t) = e^{t^2/2}}$$

With this information in hand, consider $X_1, X_2, X_3, \ldots, X_n$ identically and independently distributed random variables with cdf F. Let

$$S_n = \sum_{i=1}^{n} X_i$$

Suppose n grows without limit. What happens to S_n?

$X_1, \ldots X_n$ iid random vars

$$S_n = \sum_{i=1}^{n} X_i, \quad E(X)=0, \ Var(X)=\sigma^2$$

Set $Z_n = \dfrac{S_n}{\sigma\sqrt{n}}$. Let's look at the mgf of Z_n as n gets larger and larger. Because the X_n are indep,

$$M_{S_n}(t) = [m(t)]^n$$ (p. 13, these notes)

and

$$M_{Z_n}(t) = \left[m\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$$

We can expand m in a Taylor series around zero, so

$$M(s) = M(0) + s\,M'(0) + \frac{1}{2}s^2 M''(0) + \ldots$$

$E(x)=0$, so $M'(0)=0$

$Var(x)=\sigma^2$ so $M''(0)=\sigma^2$

remember, we've considering

$$S_n = \sum_{i=1}^{n} X_i \; ; \; Z_n = \frac{S_n}{\sigma\sqrt{n}}, \; M_{Z_n}(t) = \left[M\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$$

---

from the Taylor expansion, we now have

$$M\left(\frac{t}{\sigma\sqrt{n}}\right) = \underset{m(0)}{1} + \underset{m'(0)}{0} + \frac{1}{2}\sigma^2\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + \cdots$$

So $$M_{Z_n}(t) = \left(1 + \frac{1}{2}\sigma^2 \frac{t^2}{\sigma^2 n} + \cdots\right)^n$$

$$= \left(1 * \left(\frac{t^2}{2}\right)\frac{1}{n} + \cdots\right)^n$$

Drop the higher order terms, so

$$M_{Z_n}(t) \doteq \left(1 + \left(\frac{t^2}{2}\right)\frac{1}{n}\right)^n$$

but $\lim_{n\to\infty}\left(1 + \frac{a}{n}\right)^n = e^a$

unrigorously, we have calculated that as $n \to \infty$, $M_{Z_n}(t) \to e^{t^2/2}$.

Hence the distribution of Z_n tends to the standard normal!