MTH4106 **Introduction to Statistics**

Notes 8 **Spring 2011**

Two or more random variables

Two continuous random variables

If X and Y are continuous random variables defined on the same sample space, they have a *joint probability density function* $f_{X,Y}(x,y)$ such that, if A is any region of $\mathbb{R}^2$, then

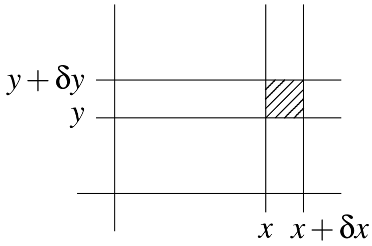
$$\mathbb{P}((X,Y) \in A) = \int\int_{(x,y)\in A} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Since probabilities are non-negative, we have $f_{X,Y}(x,y) \geqslant 0$ for all real x and y. Putting $A = \mathbb{R}^2$ gives

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1.$$

It can also be shown that if δx and δy are small and positive then

$$\mathbb{P}(x \leqslant X \leqslant x+\delta x \text{ and } y \leqslant Y \leqslant y+\delta y) \approx f_{X,Y}(x,y)\,\delta x\,\delta y.$$



The *joint cumulative distribution function* $F_{X,Y}$ is given by

$$F_{X,Y}(x,y) = \mathbb{P}(X \leqslant x \text{ and } Y \leqslant y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(t,u)\,\mathrm{d}t\,\mathrm{d}u.$$

The *marginal probability density functions* are obtained by integrating over the other variable:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y; \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x. \end{aligned}$$

If g is a real function of two variables and if X and Y have a joint continuous distribution then $\mathbb{E}(g(X,Y))$ is defined by

$$\mathbb{E}(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Theorem 8 $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

The proof is just like the proof in the discrete case, but uses integration instead of summation.

For continuous random variables, $\mathbb{P}(X = x) = \mathbb{P}(Y = y) = 0$, so we have to adapt the definition of independence.

Definition If X and Y have a joint continuous distribution then X and Y are *independent* of each other if $f_{X,Y}(x,y) = f_X(x) f_Y(y)$ for all real x and y.

If X and Y are independent of each other and $x_1 < x_2$ and $y_1 < y_2$ then

$$\begin{aligned} \mathbb{P}(x_1 \leqslant X \leqslant x_2 \text{ and } y_1 \leqslant Y \leqslant y_2) &= \int_{y_1}^{y_2}\int_{x_1}^{x_2} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{y_1}^{y_2}\int_{x_1}^{x_2} f_X(x) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{y_1}^{y_2} f_Y(y)\left[\int_{x_1}^{x_2} f_X(x)\,\mathrm{d}x\right]\mathrm{d}y \\ &= \int_{y_1}^{y_2} f_Y(y)\,\mathbb{P}(x_1 \leqslant X \leqslant x_2)\,\mathrm{d}y \\ &= \mathbb{P}(x_1 \leqslant X \leqslant x_2)\int_{y_1}^{y_2} f_Y(y)\,\mathrm{d}y \\ &= \mathbb{P}(x_1 \leqslant X \leqslant x_2)\,\mathbb{P}(y_1 \leqslant Y \leqslant y_2) \end{aligned}$$

so the events "$x_1 \leqslant X \leqslant x_2$" and "$y_1 \leqslant Y \leqslant y_2$" are independent of each other.

The following holds for joint continous random variables as well as for joint discrete random variables. Again, the proof uses integration instead of summation.

Theorem 9 If X and Y are independent then $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.

Covariance and correlation

Definition The *covariance* of X and Y is

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

This definition is exactly the same as it was for discrete random variables. So are the following two theorems and their proofs.

Theorem 10 (a) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$.

(b) $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$.

(c) If a is a constant then $\mathrm{Cov}(aX,Y) = \mathrm{Cov}(X,aY) = a\,\mathrm{Cov}(X,Y)$.

(d) If b is a constant then $\mathrm{Cov}(X,Y+b) = \mathrm{Cov}(X+b,Y) = \mathrm{Cov}(X,Y)$.

(e) $\mathrm{Cov}(X,Y+Z) = \mathrm{Cov}(X,Y) + \mathrm{Cov}(X,Z)$.

(f) If X and Y are independent then $\mathrm{Cov}(X,Y) = 0$.

Theorem 11 If a and b are constants and X and Y are random variables then

$$\mathrm{Var}(aX+bY) = a^2\,\mathrm{Var}(X) + 2ab\,\mathrm{Cov}(X,Y) + b^2\,\mathrm{Var}(Y).$$

In particular, if X and Y are independent of each other then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X-Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Definition The *correlation* between random variables X and Y is

$$\mathrm{corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}.$$

Theorem 12 $-1 \leqslant \mathrm{corr}(X,Y) \leqslant 1$.

Proof Put $\mathbb{E}(X) = \mu_X$ and $\mathbb{E}(Y) = \mu_Y$. Let t be any real number. Then

$$[(X-\mu_X) + t(Y-\mu_Y)]^2 \geqslant 0,$$

so

$$\mathbb{E}\left([(X-\mu_X) + t(Y-\mu_Y)]^2\right) \geqslant 0.$$

$$\begin{aligned}
\text{The left-hand side} &= \mathbb{E}\left[(X-\mu_X)^2+2t(X-\mu_X)(Y-\mu_Y)+t^2(Y-\mu_Y)^2\right] \\
&= \mathbb{E}\left[(X-\mu_X)^2\right]+2t\,\mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right]+t^2\,\mathbb{E}\left[(Y-\mu_Y)^2\right] \\
&= \mathrm{Var}(X)+2t\,\mathrm{Cov}(X,Y)+t^2\,\mathrm{Var}(Y) \\
&= f(t), \quad \text{say.}
\end{aligned}$$

Then $f(t)$ is a quadratic function of t which is always non-negative, so its discriminant (the part we think of as "b^2-4ac") is negative or zero: that is,

$$4\,[\mathrm{Cov}(X,Y)]^2-4\,\mathrm{Var}(X)\,\mathrm{Var}(Y)\leqslant 0.$$

The quantity $4\,\mathrm{Var}(X)\,\mathrm{Var}(Y)$ is positive, so we can divide by it without changing the sign, so

$$\frac{[\mathrm{Cov}(X,Y)]^2}{\mathrm{Var}(X)\,\mathrm{Var}(Y)}-1\leqslant 0,$$

so

$$\frac{[\mathrm{Cov}(X,Y)]^2}{\mathrm{Var}(X)\,\mathrm{Var}(Y)}\leqslant 1.$$

That is,

$$[\mathrm{corr}(X,Y)]^2\leqslant 1.$$

Taking square roots, we obtain

$$-1\leqslant \mathrm{corr}(X,Y)\leqslant 1. \quad \blacksquare$$

Theorem 13 Let X and Y be random variables such that $Y=aX+b$, where a and b are constants and $a\neq 0$. Then

$$\mathrm{corr}(X,Y)=\begin{cases}+1 & \text{if } a \text{ is positive}\\ -1 & \text{if } a \text{ is negative.}\end{cases}$$

Proof By Theorem 10, $\mathrm{Cov}(X,Y)=\mathrm{Cov}(X,aX+b)=\mathrm{Cov}(X,aX)=a\,\mathrm{Cov}(X,X)=a\,\mathrm{Var}(X)$. Also $\mathrm{Var}(Y)=\mathrm{Var}(aX+b)=\mathrm{Var}(aX)=a^2\,\mathrm{Var}(X)$, and so $\mathrm{Var}(X)\,\mathrm{Var}(Y)=a^2\,(\mathrm{Var}(X))^2$.

If $a>0$ then the positive square root of $a^2\,(\mathrm{Var}(X))^2$ is $a\,\mathrm{Var}(X)$, so

$$\mathrm{corr}(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}=\frac{a\,\mathrm{Var}(X)}{a\,\mathrm{Var}(X)}=1.$$

If $a<0$ then the positive square root of $a^2\,(\mathrm{Var}(X))^2$ is $-a\,\mathrm{Var}(X)$, so

$$\mathrm{corr}(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}=\frac{a\,\mathrm{Var}(X)}{-a\,\mathrm{Var}(X)}=-1. \quad \blacksquare$$