Stat 2911 Lecture Notes

Class 29, 2017

Uri Keich

Expectation of a function of jointly continuous RVs, Variance of the Normal and Gamma distributions, Covariance, Variance of a sum, Correlation Coefficient, Cauchy-Schwartz Inequality, Correlation of bivariate normal (Rice 4.1-4.3)

Expectation of a cont. RV

Let X be a cont. RV with pdf f_X.

- $X \in L^1$ if $\int_{-\infty}^{\infty} |x| f_X(x)\,dx < \infty$
- If $X \in L^1$ then we can define its expectation as

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

- If $X \geq 0$ $(X \leq 0)$ we can extend the definition as

$$E(X) = \int_{0}^{\infty} x f_X(x)\,dx . \qquad \left(\int_{-\infty}^{0} \text{ if } X \leq 0\right)$$

If X is a cont. RV and $g: \mathbb{R} \to \mathbb{R}$ (measurable) then with $Y = g(X)$

$$Y \in L^1 \iff \int_{\mathbb{R}} |g(x)|\, f_X(x)\,dx < \infty ,$$

and if $Y \in L^1$ (or $Y \geq 0$, or $Y \leq 0$) then

$$E(Y) = \int_{\mathbb{R}} g(x) f_X(x)\,dx .$$

<u>Cor.</u> If $\alpha, \beta \in \mathbb{R}$ and $X \in L^1$ with density f_X then $Y = \alpha X + \beta \in L^1$ and $E(Y) = \alpha E(X) + \beta$ (is Y a cont. RV?)

<u>Proof.</u>
$$\int_{-\infty}^{\infty} |\alpha x + \beta| f_X(x)\, dx \leq \int_{-\infty}^{\infty} (|\alpha x| + |\beta|) f_X(x)\, dx$$
$$\overset{?}{=} |\alpha| \underbrace{\int_{-\infty}^{\infty} |x| f_X(x)\, dx}_{< \infty \ (X \in L^1)} + |\beta| \underbrace{\int_{-\infty}^{\infty} f_X(x)\, dx}_{1}$$

$\Rightarrow Y \in L^1$ and
$$E(Y) = \int (\alpha x + \beta) f_X(x)\, dx = \dots$$
$\square$

More generally, if the random vector $\underline{X} = (X_1, \dots, X_n)$ has a joint pdf $f_{\underline{X}} : \mathbb{R}^n \to \mathbb{R}^+$, and if $g : \mathbb{R}^n \to \mathbb{R}$ (measurable) then the RV $Y = g(\underline{X})$ satisfies:

(i) $Y \in L^1 \iff \int_{\mathbb{R}^n} |g(\underline{x})| f_{\underline{X}}(x)\, dx < \infty$.

(ii) If $Y \in L^1$ (or $Y \geq 0$ or $Y \leq 0$) then
$$E(Y) = \int_{\mathbb{R}^n} g(x) f_{\underline{X}}(x)\, dx.$$

As in the discrete case, we can use this to show:

If X and Y are jointly cont. and in L^1, $\alpha, \beta \in \mathbb{R}$, then $\alpha X + \beta Y \in L^1$ and
$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y). \quad (*)$$

(replace p with f and Σ with $\int$).

We cannot prove this, but: (*) holds for *any* RVs $X, Y \in L^1$.

Also carried over from the discrete case:

If X and Y are ind. and $X, Y \in L^1$ then $XY \in L^1$ and

$$E(XY) = E(X)E(Y). \quad (p \leftrightarrow f,\ \Sigma \leftrightarrow \int)$$

For $X \in L^1$ we can define the variance of X as

$$V(X) = E[X - E(X)]^2.$$

Like in the discrete case with $p \to f$, $\Sigma \to \int$ we have

$$V(X) = E(X^2) - E^2(X). \ (*)$$

A cont. RV X is in L^k for $k \in \mathbb{N}$ if

$$\int_{-\infty}^{\infty} |x|^k f_X(x)\,dx < \infty \quad (\iff X^k \in L^1)$$

If $X \in L^2$ then $X \in L^1$ $(p \to f,\ \Sigma \to \int)$, and

$$X \in L^2 \iff V(X) < \infty \quad (\text{from } (*) \text{ as in discrete case})$$

If $\alpha, \beta \in \mathbb{R}$ and X is a RV then

$$V(\alpha X + \beta) = \alpha^2 V(X) \quad (\text{as-is})$$

If $X, Y \in L^2 \Rightarrow XY \in L^1$ (we can prove this when f_{XY} exists using $p \to f$, $\Sigma \to \int$ but the result holds more generally!)

It follows that if $X, Y \in L^2$ then $Cov(X,Y)$, initially defined as $E[(X - E(X))(Y - E(Y))]$, exists and

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

<u>Claim.</u> If $X, Y, Z \in L^2$ $\alpha, \beta \in \mathbb{R}$ then

$$Cov(\alpha X + \beta Y, Z) = \alpha Cov(X, Z) + \beta Cov(Y, Z).$$

$$Cov(\beta, Z) = 0 \quad \text{so} \quad Cov(\alpha X + \beta, Z) = \alpha Cov(X, Z)$$

(and recall the symmetry $Cov(X,Y) = Cov(Y,X)$)

<u>Proof.</u> ex.

If $X_1, \dots, X_n \in L^2$ then $\sum_1^n X_i \in L^2$ and

$$V\left(\sum_1^n X_i\right) = \sum_1^n V(X_i) + \sum_{i \neq j} Cov(X_i, X_j) \ . \quad \text{(as-is)}$$

If X_i are ind. then $V\left(\sum_1^n X_i\right) = \sum_1^n V(X_i)$ since $Cov(X_i, X_j) = 0$.

The standard deviation is defined as $\sigma_X = \sqrt{V(X)}$.
The correlation coefficient is defined as

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

provided $\sigma_X \cdot \sigma_Y > 0$.

<u>Claim</u>. Suppose $X, Y \in L^2$ with $\sigma_X \cdot \sigma_Y > 0$. Then
$$-1 \le \rho_{XY} \le 1.$$

<u>Proof</u>. Define $g: \mathbb{R} \to \mathbb{R}^+$ as $g(t) = V(t \cdot X + Y)$.

$$0 \overset{?}{\le} g(t) = V(tX) + V(Y) + 2\,Cov(tX, Y)$$
$$= t^2 \underbrace{V(X)}_{a} + \underbrace{V(Y)}_{c} + t \cdot \underbrace{2\,Cov(X,Y)}_{b}$$

$g(t) = at^2 + b \cdot t + c$ is a parabola in t $(a \ne 0)$ and since it is ≥ 0 its discriminant $b^2 - 4ac \le 0$.

$$\Rightarrow 4\,Cov^2(X,Y) \le 4\,V(X)V(Y)$$
$$\Rightarrow \rho_{XY}^2 \le 1. \qquad \square$$

Moreover,
$$|\rho_{XY}| = 1 \iff b^2 - 4ac = 0 \qquad (ac \ne 0)$$
$$\iff \exists t \text{ s.t. } 0 = g(t)$$
$$\iff \exists t \text{ s.t. } V(tX + Y) = 0$$
$$\overset{\text{exercise}}{\iff} \exists c \in \mathbb{R} \text{ s.t. } P(tX + Y = c) = 1.$$

Therefore, $|\rho_{XY}| < 1$ unless X and Y are <u>linearly dependent</u>.

Note: if X and $Y \in L^2$ and ind. then $Cov(X,Y) = 0$ and therefore $\rho_{XY} = 0$.

Examples

1) $Z \sim N(0,1)$. We saw that $E(Z)=0$.

$$E(Z^2)=\int_{-\infty}^{\infty} \underbrace{z^2}_{z \cdot z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz$$

$$= -z \cdot \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\Big]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz$$

$$= 0+1$$

$$\Rightarrow V(Z)=E(Z^2)-E^2(Z) = \quad .$$

2) If $X=\sigma Z+\mu$ where $Z\sim N(0,1)$, $\sigma>0$ and $\mu\in\mathbb{R}$, then $X\sim N(\mu,\sigma^2)$ and $f_X(x)=\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

$$E(X)=E(\sigma Z+\mu)$$
$$=\sigma E(Z)+\mu$$
$$=\mu$$

$$V(X)=V(\sigma Z+\mu)$$
$$=\sigma^2 V(Z)$$
$$=\sigma^2$$

location parameter $=\mu=E(X)$

scale parameter $=\sigma=\sigma_X$

3) $X \sim \Gamma(\alpha, \lambda)$ $\alpha, \lambda > 0$. We saw that $E(X) = \frac{\alpha}{\lambda}$.

$$E(X^2) = \int_0^\infty x^2 \frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x} dx$$

$$= \frac{1}{\lambda^2 \Gamma(\alpha)} \int_0^\infty (\lambda x)^{\alpha+1} e^{-\lambda x} \lambda dx$$

$[\lambda x = t]$

$$= \frac{1}{\lambda^2 \Gamma(\alpha)} \int_0^\infty t^{\alpha+1} e^{-t} dt$$

$$= \frac{\Gamma(\alpha+2)}{\lambda^2 \Gamma(\alpha)}$$

$$= \frac{(\alpha+1)\alpha \cdot \Gamma(\alpha)}{\lambda^2 \Gamma(\alpha)}$$

$$= \frac{\alpha(\alpha+1)}{\lambda^2}.$$

$$\Rightarrow \quad V(X) = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2$$

$$= \frac{\alpha}{\lambda^2}.$$

In particular, if $X \sim \exp(\lambda)$ then

$$E(X) = \frac{1}{\lambda}, \quad V(X) = \frac{1}{\lambda^2}.$$

If $X \sim \chi^2_n \equiv \Gamma(n/2, 1/2)$ then

$$E(X) = n, \quad V(X) = 2n.$$

4) Bivariate standard normal

$$f_{ZW}(z,w) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(z^2-2\rho z w+w^2)},$$

where $\rho \in (-1,1)$ is a parameter.

We showed that $Z, W \sim N(0,1)$

$Z|W=w \sim N(\rho w, 1-\rho^2)$

$$E(ZW) = \iint_{\mathbb{R}^2} z w f_{ZW}(z,w)\, dz\, dw$$

$$= \iint_{\mathbb{R}^2} z w f_W(w) f_{Z|W}(z|w)\, dz\, dw$$

$$= \int_w w f_W(w) \underbrace{\int_z z f_{Z|W}(z|w)\, dz}_{\rho w}\, dw$$

$$= \int \rho w^2 f_W(w)\, dw$$

$$= \rho E(W^2)$$

$$= \rho$$

$$\Rightarrow \mathrm{Cov}(Z,W) = E(ZW) - E(Z)E(W) =$$

$$\Rightarrow \rho_{ZW} = \frac{\mathrm{Cov}(Z,W)}{\sigma_Z \sigma_W} =$$

Cor. If Z and W have a standard bivariate normal dist. then Z and W are ind. iff they are uncorrelated.

Proof. For $Z, W \in L^2$ independence $\Rightarrow$ uncorrelated ($\rho_{ZW}=0$) always holds.
We showed that Z and W are ind. iff $\rho = 0$, and now we know $\rho = \rho_{ZW}$.

Recall the general bivariate normal density

$$f_{XY}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}\,\sigma_x\sigma_y} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + 2\rho\frac{x-\mu_x}{\sigma_x}\frac{y-\mu_y}{\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]}$$

where $\rho \in (-1,1)$, $\sigma_x, \sigma_y > 0$, $\mu_x, \mu_y \in \mathbb{R}$ was defined through transforming the standard bivariate normal Z, W with $\rho_{ZW} = \rho$ using: $X = \sigma_x Z + \mu_x$, $Y = \sigma_y W + \mu_y$
So $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$.

What is ρ_{XY}?

$$\begin{aligned} Cov(X,Y) &= Cov(\sigma_x Z + \mu_x, \sigma_y W + \mu_y) \\ &= \sigma_x\sigma_y \, Cov(Z,W) \\ &= \rho\,\sigma_x\sigma_y \end{aligned}$$

$$\Rightarrow \quad \rho_{xy} = \frac{Cov(X,Y)}{\sigma_x \sigma_y} =$$

<u>Cor</u> If X and Y are bivariate normal then they are ind. iff they are uncorrelated.

<u>Proof</u>. If X and Y are ind. $Cov(X,Y)=0$ so $\rho_{xy}=0$.

Conversely, $\rho = \rho_{xy} = 0$ and it is clear that f_{XY} factors into a product of $h(x)g(y)$.

Alternatively, $\rho = 0 \Rightarrow Z$ and W are ind., but $X = \varphi(Z)$ and $Y = \psi(W)$ so X and Y are ind.

If X and Y are $N(0,1)$ and $\rho_{XY}=0$ does it follow that X and Y are ind.?