Mathematics for Informatics 4a

José Figueroa-O'Farrill

Lecture 13

7 March 2012

The story of the film so far...

- C.r.v.s X and Y have a **joint density** $f(x,y)$ with

$$\mathbb{P}((X,Y)\in C) = \iint_C f(x,y)\,dx\,dy$$

and a **joint distribution**

$$F(x,y) = \mathbb{P}(X \leqslant x, Y \leqslant y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv$$

with $f(x,y) = \frac{\partial^2}{\partial x \partial y}F(x,y)$

- X and Y **independent** iff $f(x,y) = f_X(x) f_Y(y)$
- Geometric probability is fun! (Buffon's needle)
- We can calculate the c.d.f. and p.d.f. of $Z = g(X,Y)$
- X, Y independent: $f_{X+Y} = f_X \star f_Y$ (**convolution**)

Convolution

Definition

Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions. Their **convolution** $f \star g : \mathbb{R} \to \mathbb{R}$ is the function defined by

$$(f \star g)(z) = \int_{-\infty}^{\infty} f(x) g(z-x)\,dx$$

(provided the integral exists)

Properties of the convolution

- $f \star g = g \star f$
- $(f \star g) \star h = f \star (g \star h)$ (hence we can just write $f \star g \star h$)
- $f \star g$ is "smoother" than f or g

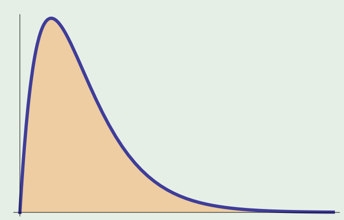
Example (Convolution of exponential variables)

- Let X and Y be independent exponentially distributed with parameter λ:

$$f_X(x) = \lambda e^{-\lambda x} \qquad f_Y(y) = \lambda e^{-\lambda y}$$

- The joint density is $f(x,y) = \lambda^2 e^{-\lambda(x+y)}$ for $x, y \geqslant 0$
- Then $Z = X + Y$ has p.d.f. given by a "gamma" distribution

$$\begin{aligned} f_Z(z) &= \int_0^{\infty} f_X(x) f_Y(z-x)\,dx \\ &= \int_0^{z} \lambda^2 e^{-\lambda x} e^{-\lambda(z-x)}\,dx \\ &= \lambda^2 z e^{-\lambda z} \end{aligned}$$



Example (Independent standard normal random variables)

- X, Y: independent, standard normally distributed. Their sum $Z = X + Y$ has p.d.f.

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-x^2/2} e^{-(z-x)^2/2} dx \\ &= \frac{e^{-z^2/4}}{2\pi} \int_{-\infty}^{\infty} e^{-(x-z/2)^2} dx && \text{(complete the square)} \\ &= \frac{e^{-z^2/4}}{2\pi} \int_{-\infty}^{\infty} e^{-u^2} du && (u = x - \tfrac{1}{2}z) \\ &= \frac{1}{2\sqrt{\pi}} e^{-z^2/4} \end{aligned}$$

  so it is normally distributed with zero mean and variance 2.
- More generally, if X has mean μ_X and variance σ_X^2 and Y has mean μ_Y and variance σ_Y^2, Z is **normally** distributed with mean $\mu_X + \mu_Y$ and variance $\sigma_X^2 + \sigma_Y^2$

Expectations of functions of random variables

- Let X and Y be c.r.v.s with joint density $f(x, y)$
- Let $Z = g(X, Y)$ for some $g : \mathbb{R}^2 \to \mathbb{R}$
- The **expectation value** of Z is defined by

$$\mathbb{E}(Z) = \iint g(x, y) f(x, y) dx\, dy$$

(provided the integral exists)

- We already saw that

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

  even if X and Y are not independent

Example (Normally distributed darts)

A dart hits a plane target at the point with coordinates (X, Y) where X and Y have joint density

$$f(x, y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

Let $R = \sqrt{X^2 + Y^2}$ be the distance from the bullseye. *What is* $\mathbb{E}(R)$*?*

$$\begin{aligned} \mathbb{E}(R) &= \iint \frac{1}{2\pi} r e^{-r^2/2} r dr\, d\theta \\ &= \int_0^{\infty} r^2 e^{-r^2/2} dr \\ &= \tfrac{1}{2} \int_{-\infty}^{\infty} r^2 e^{-r^2/2} dr \\ &= \sqrt{\tfrac{\pi}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} r^2 e^{-r^2/2} dr = \sqrt{\tfrac{\pi}{2}} \end{aligned}$$

Example (Normally distributed darts — continued)

What is $\mathbb{E}(R^2)$*?*

$$\mathbb{E}(R^2) = \mathbb{E}(X^2 + Y^2) = \mathbb{E}(X^2) + \mathbb{E}(Y^2) = 1 + 1 = 2$$

where we used

- linearity of $\mathbb{E}$, and
- the fact that $\mathbb{E}(X^2) = \mathrm{Var}(X) = 1$ and similarly for Y
- This shows that

$$\mathrm{Var}(R) = \mathbb{E}(R^2) - \mathbb{E}(R)^2 = 2 - \tfrac{\pi}{2}\ .$$

Independent random variables I

Theorem

Let X, Y be independent continuous random variables. Then

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Proof.

$$\begin{aligned}
\mathbb{E}(XY) &= \iint xyf(x,y)\,dx\,dy \\
&= \iint xyf_X(x)f_Y(y)\,dx\,dy && \text{(independence)} \\
&= \left(\int xf_X(x)\,dx\right)\left(\int yf_Y(y)\,dy\right) \\
&= \mathbb{E}(X)\mathbb{E}(Y)
\end{aligned}$$

□

Independent random variables II

As with discrete random variables, we have the following

Corollary

Let X, Y be independent continuous random variables. Then

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$$

Definition

The **covariance** and **correlation** of X and Y are

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

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\,\text{Var}(Y)}}$$

Example

Consider X, Y uniformly distributed on the unit disk D, so that

$$f(x,y) = \frac{1}{\pi}$$

Then by symmetric integration,

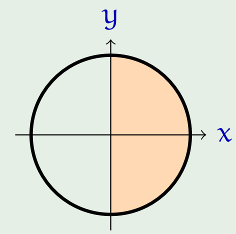
$$\mathbb{E}(XY) = \mathbb{E}(X) = \mathbb{E}(Y) = 0 \qquad \Longrightarrow \qquad \text{Cov}(X,Y) = 0$$

Therefore X, Y are uncorrelated but not independent.

Example (Continued)

On the other hand, $U = |X|$ and $V = |Y|$ are correlated.

$$\begin{aligned}
\mathbb{E}(U) &= \iint_D |x|\frac{1}{\pi}\,dx\,dy \\
&= \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_0^1 r^2\cos\theta\,dr\,d\theta \\
&= \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos\theta\,d\theta\int_0^1 r^2\,dr \\
&= \frac{2}{\pi}\times 2\times\frac{1}{3} \\
&= \frac{4}{3\pi}
\end{aligned}$$

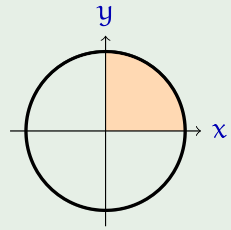


And by symmetry, also $\mathbb{E}(V) = \frac{4}{3\pi}$.

Example (Continued)

Finally,

$$\mathbb{E}(UV) = \iint_D |xy| \frac{1}{\pi} dx\, dy$$
$$= \frac{4}{\pi} \int_0^{\frac{\pi}{2}} \int_0^1 r^3 \sin\theta \cos\theta \, dr\, d\theta$$
$$= \frac{4}{\pi} \int_0^{\frac{\pi}{2}} \sin\theta\cos\theta \, d\theta \int_0^1 r^3 dr$$
$$= \frac{4}{\pi} \times \frac{1}{2} \times \frac{1}{4} = \frac{1}{2\pi}$$



Hence

$$\mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V) = \frac{1}{2\pi} - \frac{16}{9\pi^2} = \frac{9\pi - 32}{18\pi^2} < 0$$

Moment generating function of a sum

Let X, Y be independent continuous random variables and let $Z = X + Y$. Then

$$M_Z(t) = \mathbb{E}(e^{tZ})$$
$$= \int e^{tz} f_Z(z) dz$$
$$= \int e^{tz} \int f_X(x) f_Y(z-x) dx\, dz$$
$$= \iint e^{t(z-x)} e^{tx} f_X(x) f_Y(z-x) dx\, dz$$
$$= \int e^{tx} f_X(x) dx \int e^{ty} f_Y(y) dy \qquad (y = z - x)$$
$$= M_X(t) M_Y(t)$$

Markov's inequality

Theorem (Markov's inequality)

Let X be a c.r.v. Then for all $\varepsilon > 0$

$$\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \frac{\mathbb{E}(|X|)}{\varepsilon} .$$

Proof.

$$\mathbb{E}(|X|) = \int_{-\infty}^{\infty} |x| f(x) dx$$
$$= \int_{-\infty}^{-\varepsilon} |x| f(x) dx + \int_{-\varepsilon}^{\varepsilon} |x| f(x) dx + \int_{\varepsilon}^{\infty} |x| f(x) dx$$
$$\geqslant \varepsilon \int_{-\infty}^{-\varepsilon} f(x) dx + \varepsilon \int_{\varepsilon}^{\infty} f(x) dx = \varepsilon \mathbb{P}(|X| \geqslant \varepsilon)$$

□

Chebyshev's inequality

Theorem (Chebyshev's inequality)

Let X be a c.r.v. with finite mean and variance. Then

$$\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \frac{\mathbb{E}(X^2)}{\varepsilon^2} \qquad \text{for all } \varepsilon > 0$$

Proof.

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$
$$= \int_{-\infty}^{-\varepsilon} x^2 f(x) dx + \int_{-\varepsilon}^{\varepsilon} x^2 f(x) dx + \int_{\varepsilon}^{\infty} x^2 f(x) dx$$
$$\geqslant \varepsilon^2 \int_{-\infty}^{-\varepsilon} f(x) dx + \varepsilon^2 \int_{\varepsilon}^{\infty} f(x) dx = \varepsilon^2 \mathbb{P}(|X| \geqslant \varepsilon)$$

□

Two corollaries of Chebyshev's inequality

Corollary

Let X *be a c.r.v. with mean* μ *and variance* σ^2*. Then for any* $\varepsilon > 0$,

$$\mathbb{P}(|X-\mu| \geqslant \varepsilon) \leqslant \frac{\sigma^2}{\varepsilon^2}$$

Corollary (The (weak) law of large numbers)

Let $X_1, X_2, \ldots$ *be i.i.d. continuous random variables with mean* μ *and variance* σ^2 *and let* $Z_n = \frac{1}{n}(X_1 + \cdots + X_n)$*. Then*

$$\forall \varepsilon > 0 \qquad \mathbb{P}(|Z_n - \mu| < \varepsilon) \to 1 \quad \text{as } n \to \infty$$

The Chernoff bound

Corollary

Let X *be a c.r.v. with moment generating function* $M_X(t)$*. Then for any* $t > 0$,

$$\mathbb{P}(X \geqslant \alpha) \leqslant e^{-t\alpha} M_X(t)$$

Proof.

$$\mathbb{P}(X \geqslant \alpha) = \mathbb{P}(\tfrac{tX}{2} \geqslant \tfrac{t\alpha}{2}) = \mathbb{P}(e^{tX/2} \geqslant e^{t\alpha/2})$$

and by Chebyshev's inequality for $e^{tX/2}$,

$$\mathbb{P}(e^{tX/2} \geqslant e^{t\alpha/2}) \leqslant \frac{\mathbb{E}(e^{tX})}{e^{t\alpha}} = e^{-t\alpha} M_X(t) \,.$$

□

Waiting times and the exponential distribution

If "rare" and "isolated" events can occur at random in the time interval $[0, t]$, then the number of events $N(t)$ in that time interval can be approximated by a Poisson distribution

$$\mathbb{P}(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \,.$$

Let us start at $t = 0$ and let X be the time of the first event; that is, the **waiting time**. Clearly, $X > t$ if and only if $N(t) = 0$, whence

$$\mathbb{P}(X > t) = \mathbb{P}(N(t) = 0) = e^{-\lambda t} \implies \mathbb{P}(X \leqslant t) = 1 - e^{-\lambda t}$$

and differentiating,

$$f_X(t) = \lambda e^{-\lambda t}$$

whence X is exponentially distributed.

Example (Radioactivity)

The number of radioactive decays in $[0, t]$ is approximated by a Poisson distribution, so decay times are exponentially distributed. The time $t_{1/2}$ in which one half of the particles have decayed is called the **half-life**. It is a sensible concept because of the "lack of memory" of the exponential distribution.
How are the half-life and the parameter in the exponential distribution related? By definition, $\mathbb{P}(X \leqslant t_{1/2}) = \frac{1}{2}$, whence

$$e^{-\lambda t_{1/2}} = \tfrac{1}{2} \implies \lambda = \frac{\log 2}{t_{1/2}}$$

The mean of the exponential distribution: $\frac{1}{\lambda} = t_{1/2}/\log 2$ is called the **mean lifetime**.
e.g.,$t_{1/2}(^{235}\mathrm{U}) \approx 700 \times 10^6$ yrs; $t_{1/2}(^{14}\mathrm{C}) = 5{,}730$ yrs; $t_{1/2}(^{137}\mathrm{Cs}) \approx 30$ yrs

Summary

- X, Y independent random variables and $Z = X + Y$: $f_Z = f_X \star f_Y$, where $\star$ is the **convolution**
- X, Y with joint density $f(x, y)$ and $Z = g(X, Y)$:

$$\mathbb{E}(Z) = \iint g(x, y) f(x, y) dx\, dy$$

- X, Y independent:
  - $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$
  - $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$
  - $M_{X+Y}(t) = M_X(t) M_Y(t)$, where $M_X(t) = \mathbb{E}(e^{tX})$
- We defined **covariance** and **correlation** of two r.v.s
- Proved **Markov's** and **Chebyshev's** inequalities
- Proved the **(weak) law of large numbers** and the **Chernoff bound**
- Waiting times of Poisson processes are exponentially distributed