Properties of the expectation

Ignacio Cascos

2018

Outline

Introduction

We will be using the *law of iterated expectations* and the *law of conditional variances* to compute the expectation and variance of the sum of a random number of independent random variables and the expecation and variance of a mixture. Before that, we recall the formulas of the expectation and variance of a linear combination of random variables.

The second part of the session is devoted to the *moments* of a random variable.

5.1 Expectation and variance of a linear combination of random variables

We recall from *Session 4* that given d random variables $X_1, X_2, \ldots, X_d$ and real numbers $a_1, a_2, \ldots, a_d$, then

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^d a_i X_i\right] &= \sum_{i=1}^d a_i \mathbb{E}[X_i] \\
\mathrm{Var}\left[\sum_{i=1}^d a_i X_i\right] &= \sum_{i=1}^d \sum_{j=1}^d a_i a_j \mathrm{Cov}[X_i, X_j] \\
&= \sum_{i=1}^d a_i^2 \mathrm{Var}[X_i] + 2\sum_{i<j} a_i a_j \mathrm{Cov}[X_i, X_j]
\end{aligned}$$

5.2 Conditional expectation

For any X, Y random variables in the same probability space, the **conditional expectation** of X given that Y assumes value y, written as $\mathbb{E}[X|Y=y]$, is a number which is computed as

- X **discrete** $\mathbb{E}[X|Y=y] = \sum_x x p_{X|Y}(x|y)$;
- X **continuous** $\mathbb{E}[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$.

Nevertheless, $\mathbb{E}[X|Y]$ is a **random variable** that depends on Y (it is a function of random variable Y).

Law of iterated expectations

$$\mathbb{E}[X] = \mathbb{E}\left[\mathbb{E}[X|Y]\right]$$

- **Discrete random variables**

$$\mathbb{E}\left[\mathbb{E}[X|Y]\right] = \sum_y \sum_x x p_{X|Y}(x|y) p_Y(y) = \sum_x \sum_y x p_{X,Y}(x,y) = \sum_x x p_X(x) = \mathbb{E}[X].$$

- **Continuous random variables**

$$\mathbb{E}\left[\mathbb{E}[X|Y]\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X|Y}(x|y) f_Y(y) dx dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y) dy dx = \int_{-\infty}^{\infty} x f_X(x) dx = \mathbb{E}[X].$$

Conditional expectation

Mixture distribution

- If $X \sim F(x) = \sum_{i\in I} p_i F_i(x)$ and $X_i \sim F_i(x)$, that is, for some (discrete) r.v. Y, it holds $X|Y = i \sim F_i(x)$ and then $\mathbb{E}[X] = \sum_{i\in I} p_i \mathbb{E}[X|Y = i] = \sum_{i\in I} p_i \mathbb{E}[X_i]$.

Example. If $X_i \sim \mathrm{N}(\mu_i, \sigma_i)$, then $\mathbb{E}[X] = \sum_{i\in I} p_i \mathbb{E}[X_i] = \sum_{i\in I} p_i \mu_i$.

- If $X \sim F(x) = \int_A \omega(a) F_a(x) da$ and $X_a \sim F_a(x)$, that is, for some (continuous) r.v. Y, it holds $X|Y = a \sim F_a(x)$, then $\mathbb{E}[X] = \int_A \omega(a) \mathbb{E}[X_a] da$.

Example. If $X \sim \mathrm{N}(Y, \sigma)$ with $Y \sim \mathrm{U}(0,1)$, then $\mathbb{E}[X] = \int_0^1 \mathbb{E}[X|Y = y] dy = \int_0^1 y dy = 1/2$.

Conditional expectation

Sum of a random number of independent random variables

Consider $X_1, X_2, \ldots$ independent random variables with the distribution of X and N a random natural number independent of $X_1, X_2, \ldots$, then

$$\mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{N} X_i \middle| N\right]\right] = \mathbb{E}\left[N\mathbb{E}[X]\right] = \mathbb{E}[N]\mathbb{E}[X].$$

Example. We play 10 times a game. Each time we play the probability we win is 0.5 and the associated monetary prize is $\mathrm{N}(6,1)$ at each game we win, 0 when we loose. The number of victories is $N \sim \mathrm{B}(n = 10, p = 1/2)$ and the prize at our i-th win is $X_i \sim \mathrm{N}(6,1)$. Our final earnings will be $Y = \sum_{i=1}^{N} X_i$ with mean

$$\mathbb{E}[Y] = \mathbb{E}[N]\mathbb{E}[X] = 5 \times 6 = 30.$$

5.3 Conditional variance

For any X, Y random variables in the same probability space, the **conditional variance** of X given that Y assumes value y, written as $\mathrm{Var}[X|Y=y]$, is a number. We can think of it as $g(y)$.

As a function of random variable Y, the expression $\mathrm{Var}[X|Y]$, which could be written as $g(Y)$, is a **random variable** that depends on Y.

Law of conditional variances

$$\mathrm{Var}[X] = \mathbb{E}[\mathrm{Var}[X|Y]] + \mathrm{Var}[\mathbb{E}[X|Y]]$$

$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}[(X - \mathbb{E}[X|Y] + \mathbb{E}[X|Y] - \mathbb{E}[X])^2] \\ &= \mathbb{E}[(X - \mathbb{E}[X|Y])^2] + \mathbb{E}[(\mathbb{E}[X|Y] - \mathbb{E}[X])^2] + 0 \\ &= \mathbb{E}\left[\mathbb{E}[(X - \mathbb{E}[X|Y])^2]|Y\right] + \mathrm{Var}\left[\mathbb{E}[X|Y]\right] \\ &= \mathbb{E}\left[\mathrm{Var}[X|Y]\right] + \mathrm{Var}\left[\mathbb{E}[X|Y]\right].\end{aligned}$$

Conditional variance

Mixture distribution (discrete)

Assume $X \sim F(x) = \sum_{i\in I} p_i F_i(x)$ and $X_i \sim F_i(x)$. This means that for some (discrete) r.v. Y it holds $X|Y = i \sim F_i(x)$ and then

$$\mathrm{Var}[X] = \sum_{i\in I} p_i \mathrm{Var}[X_i] + \sum_{i\in I} p_i (\mathbb{E}[X_i] - \mathbb{E}[X])^2 = \sum_{i\in I} p_i \mathbb{E}[X_i^2] - \mathbb{E}[X]^2.$$

Example. If $X_i \sim \mathrm{N}(\mu_i, \sigma_i)$, then $\mathrm{Var}[X] = \sum_{i\in I} p_i(\mu_i^2 + \sigma_i^2) - \mu^2$.

Conditional variance

Mixture distribution (continuous)

Assume $X \sim F(x) = \int_A \omega(a) F_a(x) da$ and $X_a \sim F_a(x)$. This means that for some (continuous) r.v. Y, it holds $X|Y = a \sim F_a(x)$, then

$$\mathrm{Var}[X] = \int_A \omega(a)\mathrm{Var}[X_a]da + \int_A \omega(a)\left(\mathbb{E}[X_a] - \mathbb{E}[X]\right)^2 da = \int_A \omega(a)\mathbb{E}[X_a^2]da - \mathbb{E}[X]^2.$$

Example. If $X \sim \mathrm{N}(Y, \sigma)$ with $Y \sim \mathrm{U}(0,1)$, then $\mathrm{Var}[X] = \int_0^1 (\sigma^2 + y^2)dy - (1/2)^2 = \sigma^2 + 1/12$.

Conditional variance

Sum of a random number of independent random variables

Consider $X_1, X_2, \ldots$ independent random variables with the distribution of X and N a random natural number independent of $X_1, X_2, \ldots$, then

$$\begin{aligned}\mathrm{Var}\left[\sum_{i=1}^{N} X_i\right] &= \mathbb{E}\left[\mathrm{Var}\left[\sum_{i=1}^{N} X_i \middle| N\right]\right] + \mathrm{Var}\left[\mathbb{E}\left[\sum_{i=1}^{N} X_i \middle| N\right]\right] \\ &= \mathbb{E}[N\mathrm{Var}[X]] + \mathrm{Var}[N\mathbb{E}[X]] \\ &= \mathbb{E}[N]\mathrm{Var}[X] + \mathbb{E}[X]^2\mathrm{Var}[N].\end{aligned}$$

Example. We play 10 times a game… Our final earnings will be $Y = \sum_{i=1}^{N} X_i$ with variance

$$\mathrm{Var}[Y] = \mathbb{E}[N]\mathrm{Var}[X] + \mathbb{E}[X]^2\mathrm{Var}[N] = 5 \times 1 + 6^2 \times 10 \times 0.5 \times 0.5 = 95.$$

5.4 Moments of a random variable

Moments and centred moments of a random variable

If X is a random variable, and k a positive integer such that $\mathbb{E}|X|^k < \infty$, then

- the **k-th moment** of X is $\mu_k = \mathbb{E}[X^k]$;
- the **k-the centred moment** of X is $m_k = \mathbb{E}[(X - \mu_1)^k]$.

First moment (mean) location

The **first moment** of an integrable random variable is its mean (location parameter)

$$\mu_1 = \mu = \mathbb{E}[X],$$

while the **first centred moment** is 0,

$$m_1 = \mathbb{E}[X - \mu] = 0.$$

Second moment (variance) scatter

The **second moment** of a random variable with $\mathbb{E}|X|^2 < \infty$ is

$$\mu_2 = \mathbb{E}[X^2],$$

while the **second centred moment** is its variance (scatter parameter),

$$m_2 = \mathbb{E}[(X - \mu)^2] = \mathrm{Var}[X] = \sigma^2.$$

Third moment (skewness) symmetry

The *third centred moment* of a random variable can be used to obtain information about the asymmetry of its distribution

$$m_3 = \mathbb{E}[(X-\mu)^3].$$

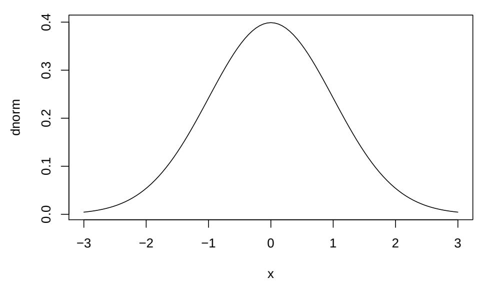
The **coefficient of skewness** is defined as

$$\text{Skew}_X = \frac{m_3}{\sigma_X^3} = \frac{\mathbb{E}[(X-\mu)^3]}{\mathbb{E}[(X-\mu)^2]^{3/2}}.$$

Third moment (skewness) symmetry

The skewness of a **symmetric distribution** is 0.

```
plot(dnorm,xlim=c(-3,3))
```



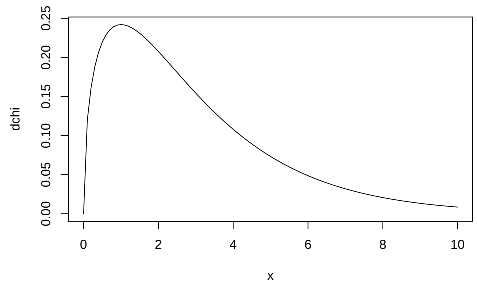

```
library(moments)
set.seed(1)
skewness(rnorm(1000))
```

```
## [1] -0.0191671
```

Third moment (skewness) positive skew

The skewness of a **right-skewed distribution** (rigth tail is longer than the left tail) is positive.

```
dchi=function(x){dchi=dchisq(x,df=3)}
plot(dchi,xlim=c(0,10))
```



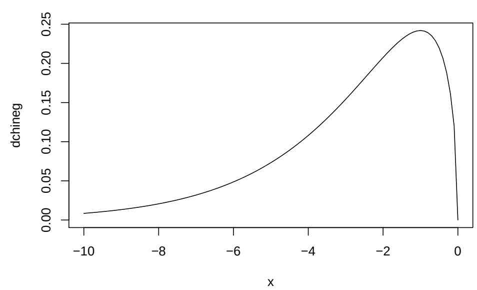
```r
set.seed(1); skewness(rchisq(1000,df=3))
```

```
## [1] 1.496328
```

Third moment (skewness) negative skew

The skewness of a **left-skewed distribution** (left tail is longer than the right tail) is negative.

```r
dchineg=function(x){dchi=dchisq(-x,df=3)}
plot(dchineg,xlim=c(-10,0))
```




```r
set.seed(1); skewness(-rchisq(1000,df=3))
```

```
## [1] -1.496328
```

Fourth moment (kurtosis) tails

The *fourth centerd moment* of a random variable can be used to obtain information about how heavy are the tails of its distribution

$$m_4 = \mathbb{E}[(X-\mu)^4].$$

The **kurtosis**

$$\text{Kurt}_X = \frac{m_4}{\sigma_X^4} = \frac{\mathbb{E}[(X-\mu)^4]}{\mathbb{E}[(X-\mu)^2]^2}$$

```r
set.seed(1)
kurtosis(rnorm(1000))
```

```
## [1] 2.998225
```

Fourth moment (kurtosis) tails

Excess kurtosis

The normal distribution is often taken as a *golden standard* and the kurtosis of a random variable is compared with that of a normal random variable by means of the **Excess kurtosis**

$$\text{EKurt}_X = \frac{m_4}{\sigma_X^4} - 3.$$

Fourth moment: zero excess kurtosis (mesokurtic)

A **mesokurtic distribution** has tails which are as heavy as those of a normal distribution and its excess kurtosis is zero.

```
set.seed(1); x.binom=rbinom(10000,size=40,prob=0.5)
hist(x.binom,probability=T)
```


```
kurtosis(x.binom)-3
```

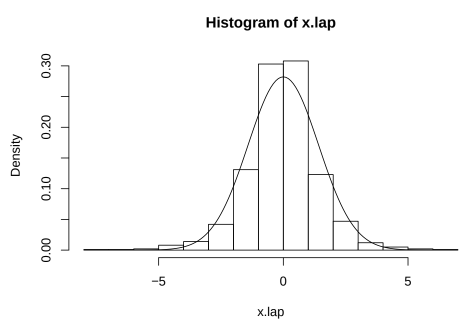
```
## [1] -0.1317299
```

Fourth moment: positive excess kurtosis (leptokurtic)

A **leptokurtic distribution** has heavier tails than a normal distribution and its excess kurtosis is positive.

```
set.seed(1); x.lap=sample(c(-1,1),1000,replace=T)*rexp(1000)
hist(x.lap,probability=T)
kurtosis(x.lap)-3
```

```
## [1] 2.371918
```

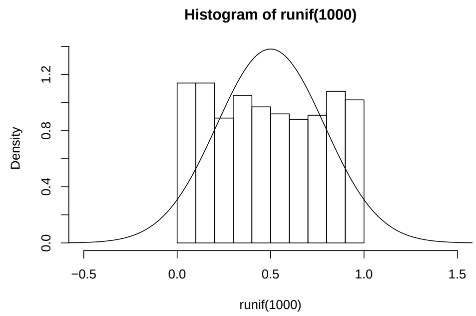


Fourth moment: negative excess kurtosis (platykurtic)

A **platykurtic distribution** has thinner tails than a normal distribution and its excess kurtosis is negative.

```
set.seed(1); x.unif=runif(1000)
hist(x.unif,probability=T)
kurtosis(x.unif)-3
```

```
## [1] -1.184201
```



5.5 The moment generating function

The **moment genearting function** of random variable X evaluated at $s \in \mathbb{R}$ is given by

$$M_X(t) = \mathbb{E}[e^{tX}]\,.$$

The **moment generating function** completely determines the distribution of the random variable X (*inversion property*).

Moment generating function

Moment generating function of some random variables

- If $Y = aX + b$, then $M_Y(t) = e^{tb} M_X(at)$
- If X and Y are independent $M_{X+Y}(t) = M_X(t) M_Y(t)$
- If $X \sim pF_{X_1} + (1-p)F_{X_2}$, then $M_X(t) = pM_{X_1}(t) + (1-p)M_{X_2}(t)$
- If $X \sim \mathrm{B}(1, p)$, then $M_X(t) = 1 - p + pe^t$
- If $X \sim \mathrm{B}(n, p)$, then $M_X(t) = (1 - p + pe^t)^n$
- If $X \sim \mathcal{P}(\lambda)$, then $M_X(t) = e^{\lambda(e^t - 1)}$
- If $X \sim \mathrm{Exp}(\lambda)$, then $M_X(t) = \frac{\lambda}{\lambda - t}$
- If $X \sim \mathrm{N}(\mu, \sigma)$, then $M_X(t) = e^{\frac{\sigma^2 t^2}{2} + \mu t}$

Moment generating function

Moment generating function and moments

The k-th derivative of the moment generating function evaluated at 0 equals the k-th moment of a random variable:

- $M_X'(t) = \mathbb{E}[Xe^{tX}]$ and $M_X'(0) = \mathbb{E}[X]$
- $M_X''(t) = \mathbb{E}[X^2 e^{tX}]$ and $M_X''(0) = \mathbb{E}[X^2]$
- $M_X^{(k)}(t) = \mathbb{E}[X^k e^{tX}]$ and $M_X^{(k)}(0) = \mathbb{E}[X^k]$