W09 Notes

Sums of random variables

01 Theory

THEOREM: Continuous PDF of a sum

Let $f_{X,Y}(x,y)$ be any joint continuous PDF.

Suppose $W = X + Y$. Then:

$$f_W(w) \;=\; \int_{-\infty}^{+\infty} f_{X,Y}(w-x, x)\,dx$$

When X and Y are *independent*, so $f_{X,Y} = f_X f_Y$, this becomes **convolution**:

$$f_W(w) \;=\; f_X * f_Y \;=\; \int_{-\infty}^{+\infty} f_X(w-x) f_Y(x)\,dx$$

- Equally valid to integrate in the y-slot: $f_W(w) \;=\; \displaystyle\int_{-\infty}^{+\infty} f_{X,Y}(x, w-x)\,dx$

Extra - Derivation of $X+Y$ PDF

The joint CDF of $X+Y$:

$$F_{X+Y}(w) = P[X+Y \leq w] \quad = \quad \iint_{x+y\leq w} f_{X,Y}(x,y)\,dx\,dy$$

Find f_{X+Y} by differentiating:

$$f_{X+Y}(w) = \frac{d}{dw} F_{X+Y}(w) \quad \gg\!\gg \quad \frac{d}{dw} \iint_{x+y\leq w} f_{X,Y}(x,y)\,dx\,dy$$

To calculate this derivative, change variables by setting $x = x$ and $s = x + y$. The Jacobian is 1, so $dx\,dy$ becomes $dx\,dw$, and we have:

$$\gg\!\gg \quad \frac{d}{dw} \int_{-\infty}^{w} \int_{-\infty}^{+\infty} f_{X,Y}(x, s-x)\,dx\,ds \quad \gg\!\gg \quad \int_{-\infty}^{+\infty} f_{X,Y}(x, w-x)\,dx$$

02 Illustration

Example - Sum of parabolic random variables

Suppose X is an RV with PDF given by:

$$f_X(x) = \begin{cases} \frac{3}{4}(1-x^2) & x \in [-1,1] \\ 0 & \text{otherwise} \end{cases}$$

Let Y be an independent copy of X. So $f_Y = f_X$, but Y is independent of X.

Find the PDF of $X+Y$.

Solution

The graph of $f_X(w-x)$ matches the graph of $f_X(x)$ except (i) flipped in a vertical mirror, (ii) shifted by w to the left.

When $w \in [-2,0]$, the integrand is nonzero only for $x \in [-1, w+1]$:

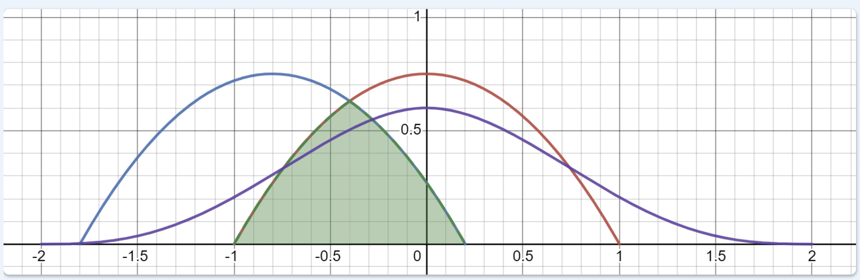
$$\begin{aligned} f_{X+Y}(w) &= \left(\frac{3}{4}\right)^2 \int_{-1}^{w+1} \left(1-(w-x)^2\right)\left(1-x^2\right) dx \\ &= \frac{9}{16}\left(\frac{w^5}{30} - \frac{2w^3}{3} - \frac{4w^2}{3} + \frac{16}{15}\right) \end{aligned}$$

When $w \in [0,+2]$, the integrand is nonzero only for $x \in [w-1, +1]$:

$$\begin{aligned} f_{X+Y}(w) &= \left(\frac{3}{4}\right)^2 \int_{w-1}^{+1} \left(1-(w-x)^2\right)\left(1-x^2\right) dx \\ &= \frac{9}{16}\left(-\frac{w^5}{30} + \frac{2w^3}{3} - \frac{4w^2}{3} + \frac{16}{15}\right) \end{aligned}$$

Final result is:

$$f_{X+Y}(w) = \begin{cases} \frac{9}{16}\left(\frac{w^5}{30} - \frac{2w^3}{3} - \frac{4w^2}{3} + \frac{16}{15}\right) & w \in [-2,0] \\ \frac{9}{16}\left(-\frac{w^5}{30} + \frac{2w^3}{3} - \frac{4w^2}{3} + \frac{16}{15}\right) & w \in [0,2] \\ 0 & \text{otherwise} \end{cases}$$

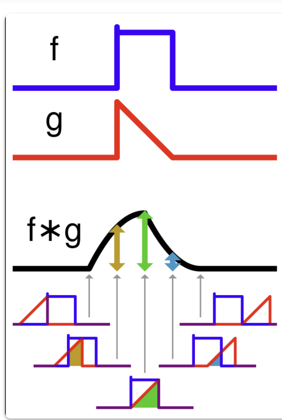


03 Theory - extra

Convolution

The **convolution** of two continuous functions $f(x)$ and $g(x)$ is defined by:

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(x-t)g(t)\, dt$$



For more example calculations, look at 9.6.1 and 9.6.2 at this page.

Applications of convolution

- Convolutional neural networks (machine learning theory: translation invariant NN, low pre-processing)
- Image processing: edge detection, blurring
- Signal processing: smoothing and interpolation estimation
- Electronics: linear translation-invariant (LTI) system response: convolution with impulse function

Extra - Convolution

Geometric meaning of convolution

Convolution does not have a neat and precise geometric meaning, but it does have an imprecise intuitive sense.

The product of two quantities tends to be large when *both* quantities are large; when one of them is small or zero, the product will be small or zero. This behavior is different from the behavior of a sum, where one summand being large is sufficient for the sum to be large. A large summand overrides a small co-summand, whereas a large factor is scaled down by a small cofactor.

The upshot is that a convolution will be large when two functions *have similar overall shape*. (Caveat: one function must be flipped in a vertical mirror before the overlay is considered.) The argument value where the convolution is largest will correspond to the horizontal offset needed to get the closest overlay of the functions.

Algebraic properties of convolution

- $f * g = g * f$
- $f * (g * h) = (f * g) * h$
- $f * (g + h) = f * g + f * h$
- $a(f * g) = (af) * g = f * (ag)$
- $(f * g)' = f' * g = f * g'$

The last of these is *not* the typical Leibniz rule for derivatives of products!

All of these properties can be checked by simple calculations with iterated integrals.

Convolution in more variables

Given $f, g : \mathbb{R}^n \to \mathbb{R}$, their convolution at $\mathbf{x}$ is defined by integrating the shifted products over the whole domain:

$$(f * g)(\mathbf{x}) = \iiint_{\mathbb{R}^n} f(\mathbf{x} - \mathbf{y}) g(\mathbf{y}) \, dy$$

04 Illustration

Exercise - Convolution practice

Suppose X is an RV with density:

$$f_X = \begin{cases} 2x & x \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

Suppose Y is uniform on $[0, 1]$.

Find the PDF of $X + Y$. Sketch the graph of this PDF.

05 Theory

Recall that in a Poisson process:

- $X \sim \text{Exp}(\lambda)$ measures continuous wait time until *one* arrival
- $X \sim \text{Erlang}(\ell, \lambda)$ measures continuous wait time until ℓ^{th} arrival

Since the wait times between arrivals are *independent*, we expect that the *sum of exponential distributions is an Erlang distribution*. This is true!

Erlang sum rule

Specify a given Bernoulli process with success probability p. Suppose that:

- $X \sim \text{Erlang}(r, \lambda)$
- $Y \sim \text{Erlang}(s, \lambda)$
- X and Y are independent

Then:

$$X + Y \quad \sim \quad \text{Erlang}(r + s, \lambda)$$

Exp plus Exp is Erlang

Recall that $\text{Erlang}(1, \lambda) \sim \text{Exp}(\lambda)$.

So we could say:

$$\text{“}\text{Exp}(\lambda) + \text{Exp}(\lambda) \ \sim \text{Erlang}(2, \lambda)\text{”}$$

And:

$$\text{“}\text{Exp}(\lambda) + \text{Erlang}(\ell, \lambda) \ \sim \text{Erlang}(\ell + 1, \lambda)\text{”}$$

06 Illustration

Example - Exp plus Exp equals Erlang

Let us verify this formula by direct calculation:

$$\text{“}\text{Exp}(\lambda) + \text{Exp}(\lambda) \ \sim \ \text{Erlang}(2, \lambda)\text{”}$$

Solution

Let $X, Y \sim \text{Exp}(\lambda)$ be independent RVs.

Therefore:

$$f_X = f_Y \quad = \quad \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Now compute the convolution:

$$
\begin{aligned}
f_{X+Y}(w) \quad &= \quad \int_{-\infty}^{+\infty} f_X(w-x) f_Y(x)\,dx \\
&\gg\!\gg \quad \int_0^w \lambda^2 e^{-\lambda(w-x)} e^{-\lambda x}\,dx \\
&\gg\!\gg \quad \lambda^2 \int_0^w e^{-\lambda w}\,dx \quad \gg\!\gg \quad \lambda^2 w e^{-\lambda w}
\end{aligned}
$$

This is the Erlang PDF:

$$
f_X(t) = \frac{\lambda^\ell}{(\ell-1)!} t^{\ell-1} e^{-\lambda t} \bigg|_{\ell=2}
$$

Exercise - Erlang induction step

By direct computation with PDFs and convolution, derive the formula:

$$
\text{“Exp}(\lambda) + \text{Erlang}(\ell,\lambda) \;\sim\; \text{Erlang}(\ell+1,\lambda)\text{”}
$$

Observation: By repeatedly applying the above formula, it follows that:

$$
\text{“}\overbrace{\text{Exp}(\lambda) + \cdots + \text{Exp}(\lambda)}^{\ell \text{ terms}} \;\sim\; \text{Erlang}(\ell,\lambda)\text{”}
$$

Expectation for two variables

07 Theory

Expectation for a function on two variables

Discrete case:

$$
E[\,g(X,Y)\,] \quad = \quad \sum_{k,\ell} g(k,\ell)\, P_{X,Y}(k,\ell) \qquad \text{(sum over possible values)}
$$

Continuous case:

$$
E[\,g(X,Y)\,] \quad = \quad \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y)\, f_{X,Y}(x,y)\,dx\,dy
$$

These formulas are *not trivial to prove*, and we omit the proofs. (Recall the technical nature of the proof we gave for $E[\,g(X)\,]$ in the discrete case.)

Expectation sum rule

Suppose X and Y are *any* two random variables on the same probability model.

Then:

$$
E[X+Y] = E[X] + E[Y]
$$

We already know that expectation is linear in a single variable: $E[aX+b] = aE[X]+b$.

Therefore this two-variable formula implies:

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

Expectation product rule: independence

Suppose that X and Y are *independent*.

Then we have:

$$E[XY] = E[X]E[Y]$$

Extra - Proof: Expectation sum rule, continuous case

Suppose f_X and f_Y give marginal PDFs for X and Y, and $f_{X,Y}$ gives their joint PDF.

Then:

$$\begin{aligned} E[X+Y] \quad & \gg\gg \quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f_{X,Y}(x,y)\,dx\,dy \\ & \gg\gg \quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf_{X,Y}\,dx\,dy + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf_{X,Y}\,dx\,dy \\ & \gg\gg \quad \int_{-\infty}^{+\infty} xf_X(x)\,dx + \int_{-\infty}^{+\infty} yf_Y(y)\,dy \\ & \gg\gg \quad E[X] + E[Y] \end{aligned}$$

Observe that this calculation relies on the formula for $E[\,g(X,Y)\,]$, specifically with $g(x,y) = x + y$.

Extra - Proof: Expectation product rule

$$\begin{aligned} E[XY] \quad & \gg\gg \quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(xy)f_{X,Y}(x,y)\,dx\,dy \\ & \gg\gg \quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(xy)f_X(x)f_Y(y)\,dx\,dy \\ & \gg\gg \quad \int_{-\infty}^{+\infty} xf_X(x)\,dx \int_{-\infty}^{+\infty} yf_Y(y)\,dy \\ & \gg\gg \quad E[X]E[Y] \end{aligned}$$

08 Illustration

$E[\,X^2 + Y\,]$ from joint PMF chart

Suppose the joint PMF of X and Y is given by this chart:

$Y \downarrow \ X \rightarrow$	1	2
-1	0.2	0.2
0	0.35	0.1
1	0.05	0.1

Define $W = X^2 + Y$. Find the expectation $E[W]$.

Solution

First compute the values of W for each pair (X, Y) in the chart:

$Y \downarrow \; X \rightarrow$	1	2
-1	0	3
0	1	4
1	2	5

Now take the sum, weighted by probabilities:

$$\begin{array}{c} 0 \cdot (0.2) + 3 \cdot (0.2) + 1 \cdot (0.35) \\ +4 \cdot (0.1) + 2 \cdot (0.05) + 5 \cdot (0.1) \end{array} \quad \gg\gg \quad 1.95 \; = \; E[W]$$

Exercise - Understanding expectation for two variables

Suppose you know *only* that $X \sim \text{Geo}(p)$ and $Y \sim \text{Bin}(n, q)$.

Which of the following can you calculate?

$$E[X+Y], \quad E[XY], \quad E[X^2+Y^2], \quad E[(X+Y)^2]$$

$E[Y]$ two ways, and $E[XY]$, from joint density

Suppose X and Y are random variables with the following joint density:

$$f_{X,Y}(x,y) = \begin{cases} \frac{3}{16}xy^2 & x, y \in [0,2] \\ 0 & \text{otherwise} \end{cases}$$

(a) Compute $E[Y]$ using two methods.

(b) Compute $E[XY]$.

Solution

(a)

(1) <u>Method One:</u> via marginal PDF $f_Y(y)$:

$$f_Y(y) \quad = \quad \int_0^2 \frac{3}{16}xy^2\,dx \quad \gg\gg \quad \begin{cases} \frac{3}{8}y^2 & y \in [0,2 \\ 0 & \text{otherwise} \end{cases}$$

Then expectation:

$$E[Y] \; = \; \int_0^2 y\, f_Y(y)\,dy \quad \gg\gg \quad \int_0^2 \frac{3}{8}y^3\,dy \quad \gg\gg \quad 3/2$$

(2) <u>Method Two:</u> directly, via two-variable formula:

$$E[Y] \; = \; \int_0^2 \int_0^2 y \cdot \frac{3}{16}xy^2\,dy\,dx \gg\gg \quad \int_0^2 \frac{3}{4}x\,dx \quad \gg\gg \quad 3/2$$

(b) Directly, via two-variable formula:

$$E[XY] \quad = \quad \int_0^2 \int_0^2 xy \cdot \frac{3}{16} xy^2 \, dy \, dx$$

$$\gg\gg \quad \int_0^2 \frac{3}{4} x^2 \, dx \quad \gg\gg \quad 2$$

Covariance and correlation

09 Theory

Write $\mu_X = E[X]$ and $\mu_Y = E[Y]$.

Observe that the random variables $X - \mu_X$ and $Y - \mu_Y$ are "centered at zero," meaning that $E[X - \mu_X] = 0 = E[Y - \mu_Y]$.

Covariance

Suppose X and Y are any two random variables on a probability model. The **covariance** of X and Y measures the *typical synchronous deviation* of X and Y from their respective means.

Then the *defining formula* for covariance of X and Y is:

$$\text{Cov}[X, Y] \;=\; E[\,(X - \mu_X)(Y - \mu_Y)\,]$$

There is also a *shorter formula*:

$$\text{Cov}[X, Y] \;=\; E[XY] - \mu_X \mu_Y$$

To derive the shorter formula, first expand the product $(X - \mu_X)(Y - \mu_Y)$ and then apply linearity.

Notice that covariance is always *symmetric*:

$$\text{Cov}[X, Y] \quad = \quad \text{Cov}[Y, X]$$

The *self* covariance equals the variance:

$$\text{Cov}[X, X] \quad = \quad \text{Var}[X]$$

The *sign* of $\text{Cov}[X, Y]$ reveals the *correlation type* between X and Y:

Correlation	Sign
Positively correlated	$\text{Cov}(X, Y) > 0$
Negatively correlated	$\text{Cov}(X, Y) < 0$
Uncorrelated	$\text{Cov}(X, Y) = 0$

Correlation coefficient

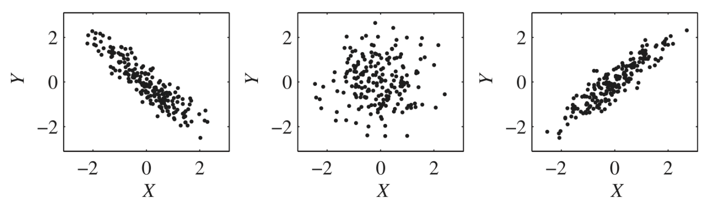
Suppose X and Y are any two random variables on a probability model.

Their **correlation coefficient** is a rescaled version of covariance that measures the *synchronicity of deviations*:

$$\rho[X, Y] \;=\; \frac{\text{Cov}[X, Y]}{\sigma_X \sigma_Y}$$

The rescaling ensures:

$$-1 \;\leq\; \rho_{X,Y} \;\leq\; +1$$



(a) $\rho_{X,Y} = -0.9$ (b) $\rho_{X,Y} = 0$ (c) $\rho_{X,Y} = 0.9$

Covariance depends on the *separate variances* of X and Y as well as their relationship.

Correlation coefficient, because we have divided out $\sigma_X\sigma_Y$, depends only on their *relationship*.

10 Illustration

Covariance from PMF chart

Suppose the joint PMF of X and Y is given by this chart:

$Y \downarrow \;\; X \rightarrow$	1	2
-1	0.2	0.2
0	0.35	0.1
1	0.05	0.1

Find $\mathrm{Cov}[X, Y]$.

Solution

We need $E[X]$ and $E[Y]$ and $E[XY]$.

$$E[X] \;=\; 1 \cdot (0.2 + 0.35 + 0.05) + 2 \cdot (0.2 + 0.1 + 0.1) \quad \gg\!\gg \quad 1.4$$

$$E[Y] \;=\; -1 \cdot (0.2 + 0.2) + 0 \cdot (0.35 + 0.1) + 1 \cdot (0.05 + 0.1)$$

$$\gg\!\gg \quad -0.25$$

$$E[XY] \;=\; -1 \cdot (0.2) - 2 \cdot (0.2) + 0 + 1 \cdot (0.05) + 2 \cdot (0.1) \quad \gg\!\gg \quad -0.35$$

Therefore:

$$\mathrm{Cov}[X, Y] \;=\; E[XY] - E[X]E[Y]$$

$$\gg\!\gg \quad -0.35 - (1.4)(-0.25) \quad \gg\!\gg \quad 0$$

11 Theory

Covariance bilinearity

Given any three random variables X, Y, and Z, we have:

$$\mathrm{Cov}[\,X + Y,\; Z\,] \;=\; \mathrm{Cov}[X, Z] + \mathrm{Cov}[Y, Z]$$

$$\mathrm{Cov}[\,Z,\; X + Y\,] \;=\; \mathrm{Cov}[Z, X] + \mathrm{Cov}[Z, Y]$$

Covariance and correlation: shift and scale

Covariance *scales with each input, and ignores shifts*:

$$\mathrm{Cov}[\,aX+b,\,Y\,] \quad = \quad a\,\mathrm{Cov}[X,Y] \quad = \quad \mathrm{Cov}[\,X,\,aY+b\,]$$

Whereas shift or scale in correlation *only affects the sign*:

$$\rho[\,aX+b,\,Y\,] \quad = \quad \mathrm{sign}(a)\,\rho[X,Y] \quad = \quad \rho[\,X,\,aY+b\,]$$

Extra - Proof of covariance bilinearity

$$\begin{aligned}
\mathrm{Cov}[X+Y,\,Z] \quad &\ggg \quad E[(X+Y-(\mu_X+\mu_Y))(Z-\mu_Z)] \\
&\ggg \quad E[(X-\mu_X+Y-\mu_Y)(Z-\mu_Z)] \\
&\ggg \quad E[(X-\mu_X)(Z-\mu_Z)]+E[(Y-\mu_Y)(Z-\mu_Z)] \\
&\ggg \quad \mathrm{Cov}[X,Z]+\mathrm{Cov}[Y,Z]
\end{aligned}$$

Extra - Proof of covariance shift and scale rule

$$\begin{aligned}
\mathrm{Cov}[aX+b,Y] \quad &\ggg \quad E[(aX+b)Y]-E[aX+b]E[Y] \\
&\ggg \quad E[aXY+bY]-aE[X]E[Y]-E[b]E[Y] \\
&\ggg \quad aE[XY]+bE[Y]-aE[X]E[Y]-bE[Y] \\
&\ggg \quad a\big(E[XY]-E[X]E[Y]\big)
\end{aligned}$$

Independence implies zero covariance

Suppose that X and Y are any two random variables on a probability model.

If X and Y are independent, then:

$$\mathrm{Cov}[X,Y]=0$$

Proof:

We know both of these:

$$E[XY]=E[X]E[Y] \qquad \text{(independence)}$$

$$\mathrm{Cov}[X,Y] \;=\; E[XY]-\mu_X\mu_Y \qquad \text{(shorter form)}$$

But $E[XY]=E[X]E[Y]=\mu_X\mu_Y$, so those terms cancel and $\mathrm{Cov}[X,Y]=0$.

Sum rule for variance

Suppose that X and Y are any two random variables on a probability space.

Then:

$$\mathrm{Var}[X+Y]=\mathrm{Var}[X]+\mathrm{Var}[Y]+2\mathrm{Cov}[X,Y]$$

When X and Y are *independent*:

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

Extra - Proof: Sum rule for variance

$$\begin{aligned}\text{Var}[X+Y] \quad &\gg\!\gg \quad E\Big[\big(X+Y-(\mu_X+\mu_Y)\big)^2\Big] \\ &\gg\!\gg \quad E\Big[\big((X-\mu_X)+(Y-\mu_Y)\big)^2\Big] \\ &\gg\!\gg \quad E\Big[(X-\mu_X)^2+(Y-\mu_Y)^2+2(X-\mu_X)(Y-\mu_Y)\Big] \\ &\gg\!\gg \quad \text{Var}[X]+\text{Var}[Y]+2\text{Cov}[X,Y]\end{aligned}$$

Extra - Proof that $-1 \leq \rho \leq +1$

(1) Create standardizations:

$$\tilde{X} \;=\; \frac{X-\mu_X}{\sigma_X}, \qquad \tilde{Y} \;=\; \frac{Y-\mu_Y}{\sigma_Y}$$

Now $\tilde{X}$ and $\tilde{Y}$ satisfy:

$$E[\tilde{X}]=0=E[\tilde{Y}] \qquad \text{and} \qquad \text{Var}[\tilde{X}]=1=\text{Var}[\tilde{Y}]$$

Observe that $\text{Var}[W]\geq 0$ for any W. *Variance can't be negative.*

---

(2) Apply the variance sum rule.

Apply to $\tilde{X}$ and $\tilde{Y}$:

$$0 \leq \text{Var}[\tilde{X}+\tilde{Y}] \;=\; \text{Var}[\tilde{X}]+\text{Var}[\tilde{Y}]+2\text{Cov}[\tilde{X},\tilde{Y}]$$

Simplify:

$$\gg\!\gg \quad 1+1+2\text{Cov}[\tilde{X},\tilde{Y}]\geq 0$$

$$\gg\!\gg \quad 1+\text{Cov}[\tilde{X},\tilde{Y}]\geq 0$$

Notice effect of standardization:

$$\text{Cov}[\tilde{X},\tilde{Y}] \quad = \quad \rho[X,Y]$$

Therefore $\rho[X,Y]\geq -1$.

---

(3) Modify and reapply variance sum rule.

Change to $\tilde{X}-\tilde{Y}$:

$$0 \leq \text{Var}[\tilde{X}-\tilde{Y}] \;=\; \text{Var}[\tilde{X}]+\text{Var}[-\tilde{Y}]+2\text{Cov}[\tilde{X},\,-\tilde{Y}]$$

Simplify:

$$\gg\!\gg \quad 1+1-2\text{Cov}[\tilde{X},\tilde{Y}]\geq 0$$

$$\gg\!\gg \quad 1-\text{Cov}[\tilde{X},\tilde{Y}]\geq 0$$

12 Illustration

Variance of sum of indicators

An urn contains 3 red balls and 2 yellow balls.

Suppose 2 balls are drawn without replacement, and X counts the number of red balls drawn.

Find $\mathrm{Var}[X]$.

Solution

Let X_1 indicate (one or zero) whether the first ball is red, and X_2 indicate whether the second ball is red, so $X = X_1 + X_2$.

Then $X_1 X_2$ indicates whether both drawn balls are red; so it is Bernoulli with success probability $\frac{3}{5}\frac{2}{4} = \frac{3}{10}$. Therefore $E[X_1 X_2] = \frac{3}{10}$.

We also have $E[X_1] = E[X_2] = \frac{3}{5}$.

The variance sum rule gives:

$$\begin{aligned} \mathrm{Var}[X] &= \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + 2\mathrm{Cov}[X_1, X_2] \\ &\ggg \quad E[X_1^2] - E[X_1]^2 + E[X_2^2] - E[X_2]^2 + 2(E[X_1 X_2] - E[X_1]E[X_2]) \\ &\ggg \quad \frac{3}{5} - \left(\frac{3}{5}\right)^2 + \frac{3}{5} - \left(\frac{3}{5}\right)^2 + 2\left(\frac{3}{10} - \frac{3}{5}\cdot\frac{3}{5}\right) \quad \ggg \quad \frac{9}{25} \end{aligned}$$

Exercise - Covariance rules

Simplify:

$$\mathrm{Cov}[\,2X + 5Y + 1,\; Z + 8W + X + 9\,]$$

Exercise - Independent variables are uncorrelated

Let X be given with possible values $\{-1, 0, +1\}$ and PMF given uniformly by $P_X(k) = 1/3$ for all three possible k. Let $Y = X^2$.

Show that X and Y are dependent but uncorrelated.

Hint: To speed the calculation, notice that $X^3 = X$.