Mathematics 50
Probability and Statistical Inference
Winter 1999

2. Multivariate Continuous Random Variables

Week 4
January 25-29

1. Distribution and density functions

Two CRVs are considered, X and Y. The *joint* cumulative (bivariate) distribution function (df) of (X, Y) is

$$F_{X,Y}(x, y) = F(x, y) = \Pr(X \leq x, Y \leq y).$$

The joint density is defined as

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

Back representation:

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y) dx dy,$$

Main Theorem of Calculus.

What is the probability $(X, Y) \in A$? Answer,

$$\Pr(A) = \int_A f(x, y) dx dy.$$

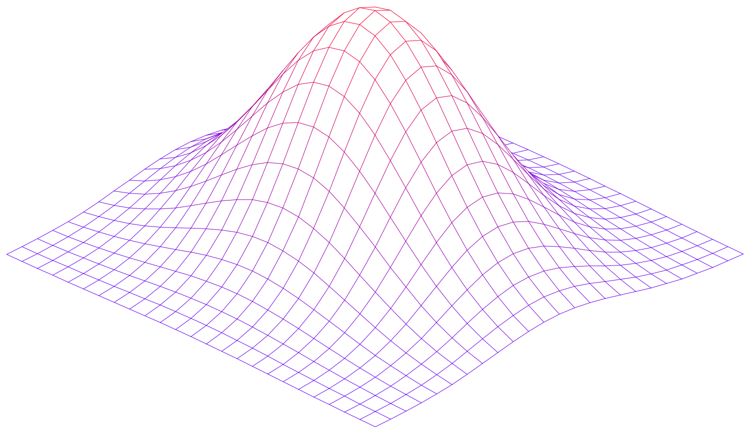
How to find $\Pr(X + Y \leq c)$?

$$\Pr(X + Y \leq c) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{c-x} f(x, y) dy \right) dx.$$

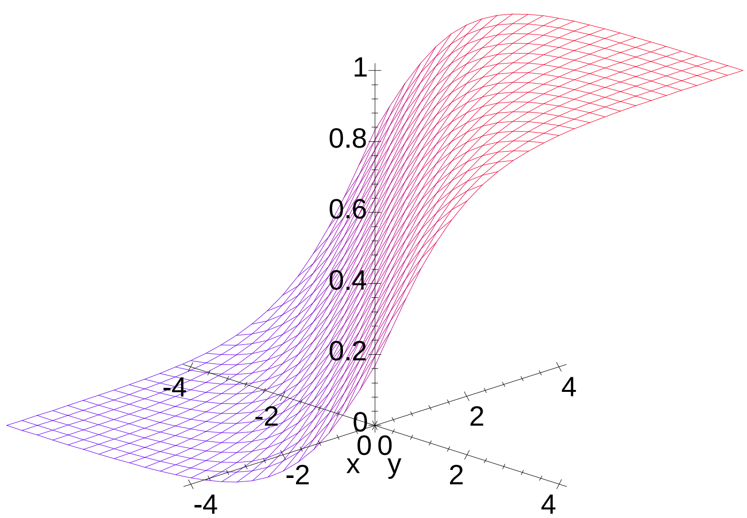
Properties of multivariate distributions:

1. $F(-\infty, -\infty) = 0, F(\infty, \infty) = 1.$
2. $F(x, y)$ is an increasing function in both arguments: if $x_1 \leq x_2$ and $y_1 \leq y_2$ then $F(x_1, y_1) \leq F(x_2, y_2)$.
3. $f(x_1, x_2) \geq 0.$
4. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1.$

To prove that a certain function is a distribution function we must show that 1 and 2 hold. To prove that a certain function is a density function we must prove that 3 and 4 hold.



Typical density function of a CRV



Distribution function of a multivaraite logistic distribution $\exp(x+y)/(1+\exp(x+y))$

Problem (Rice, p.74). The bivariate density function is defined as

$$f(x,y) = \begin{cases} c(x^2+xy) & \text{if} \quad 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \qquad \text{elsewhere} \end{cases}$$

Find c and $\Pr(X > Y)$.

Solution. We find c from the condition

$$\int_0^1 \int_0^1 f(x,y)dydx = \int_0^1 \left(\int_0^1 f(x,y)dy \right) dx = 1.$$

We first evaluate the inner integral: if x is fixed then

$$\int_0^1 (x^2 + xy)dy = x^2 + x \int_0^1 ydy = x^2 + \frac{1}{2}x.$$

Now we evaluate the outer integral

$$\int_0^1 \left(x^2 + \frac{1}{2}x \right) dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}.$$

Therefore,

$$c = \frac{12}{7}.$$

The needed probability is

$$\Pr(X > Y) = \int_0^1 \int_0^x f(x,y)dydx = \frac{12}{7} \int_0^1 \int_0^x (x^2 + xy)dydx = \frac{12}{7} \int_0^1 \left(\int_0^x (x^2 + xy)dy \right) dx.$$

Again, calculate first the inner integral (x fixed)

$$\int_0^x (x^2 + xy)dy = x^3 + x \int_0^x ydy = x^3 + \frac{1}{2}x^3 = \frac{3}{2}x^3.$$

Finally,

$$\Pr(X > Y) = \frac{12}{7} \times \frac{3}{2} \int_0^1 x^3 dx = \frac{12}{7} \times \frac{3}{2} \times \frac{1}{4} = \frac{9}{14}.$$

2. Marginal distribution

If the joint distribution is known, what is the distribution of each component (regardless of the other)? Reasonably, we define the *marginal* distribution function of X as

$$\begin{aligned} F_X(x) &= F(X \le x, -\infty < Y < \infty) \\ &= F(x, \infty) = \Pr(X \le x). \end{aligned}$$

or

$$F_X(x) = \lim_{y \to \infty} F(x,y)$$

According to the above,

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

i.e. y is integrated out.

Problem (continued). Find marginal densities of X and Y.

Solution. By definition (x fixed and y integrates out),

$$\begin{aligned} f_X(x) &= \int_0^1 f(x,y)dy = \frac{12}{7}\int_0^1 (x^2+xy)dy \\ &= \frac{12}{7}x\left(x+\frac{1}{2}\right), \text{for } 0<x<1. \end{aligned}$$

Analogously (y fixed and x integrates out),

$$\begin{aligned} f_Y(y) &= \int_0^1 f(x,y)dx = \frac{12}{7}\int_0^1 (x^2+xy)dx = \frac{12}{7}\int_0^1 x^2dx + \frac{12}{7}y\int_0^1 xdx \\ &= \frac{12}{7}\times\frac{1}{3}+\frac{12}{7}y\times\frac{1}{2} = \frac{2}{7}(2+3y) \text{ for } 0<y<1. \end{aligned}$$

3. Independence

is an important concept in probability and statistics.

Recall, two events A and B are independent if and only of

$$\Pr(A\cap B) = \Pr(A)\times\Pr(B).$$

Similarly, we say that X and Y are independent if

$$\begin{aligned} F(x,y) &= \Pr(X\le x, Y\le y) = \Pr(X\le x)\times\Pr(Y\le y) \\ &= F_X(x)F_Y(y). \end{aligned} \tag{3.1}$$

Therefore, two RVs are independent if the joint d.f. is the product of the marginal d.f.

Take derivative and find the condition on independence expressed in densities

$$f(x,y) = f_X(x)f_Y(y). \tag{3.2}$$

The conditions (3.1) and (3.2) are equivalent.

Problem (continued). Are X and Y independent?

Solution. X and Y are independent iff

$$f(x,y) = f_X(x)f_Y(y).$$

We have

$$\begin{aligned} f_X(x) &= \frac{12}{7}x\left(x+\frac{1}{2}\right), \\ f_Y(y) &= \frac{2}{7}(2+3y), \\ f(x,y) &= \frac{12}{7}x(x+y). \end{aligned}$$

We see that

$$f(x,y) \neq f_X(x)\times f_Y(y).$$

Answer:X are Y **dependent.**

The definition with two RVs is easy to extend to the case of several RVs. A very important case is when n RVs are *independent and identically distributed* (iid).

We write $X_1, X_2, ..., X_n$ are iid.

Example. Uniform distribution on the plane (bivariate uniform distribution).
A pair of CRVs has multivariate uniform distribution on $(0, a) \times (0, a)$ if the d.f. is specified as

$$F(x, y) = \frac{1}{a^2} \times \begin{cases} xy \text{ if } 0 < x < a, 0 < y < a \\ 0 \text{ otherwise.} \end{cases}$$

Find the joint density, marginal distributions and density functions. Are X and Y independent?
Solution. The joint density is

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{1}{a^2} = const.$$

Marginal distribution function of X is

$$F_X(x) = \lim_{y \to \infty} F(x, y) = F(x, y = a) = \frac{1}{a^2} xa = \frac{x}{a}.$$

Marginal distribution function of Y

$$F_Y(y) = \lim_{x \to \infty} F(x, y) = F(x = a, y) = \frac{1}{a^2} ay = \frac{y}{a}.$$

Both have uniform distribution on $(0, a)$. They are independent because $F(x, y) = F_X(x) F_Y(y)$:

$$\frac{1}{a^2} xy = \frac{x}{a} \times \frac{y}{a}.$$

Problem. Let $f(x)$ be a density. A function of two arguments is defined as

$$w(x, y) = f(x + y)$$

Is w a joint density function?
Solution. The marginal density is

$$f_X(x) = \int_{-\infty}^{\infty} f(x + y) dy.$$

But

$$\int_{-\infty}^{\infty} f(x + y) dy = \int_{-\infty}^{\infty} f(u) du = 1.$$

So we obtain, the marginal density is

$$f_X(x) = 1 \text{ for all } x.$$

Therefore, $f_X(x)$ cannot be a density, and w is not a joint density function.

4. Conditional distribution

Given joint distribution for the pair RVs (X, Y), what is the conditional distribution of Y when X fixed at the value $X = x$? Recall the conditional probability is defined as

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

We say "conditional probability of A given B"

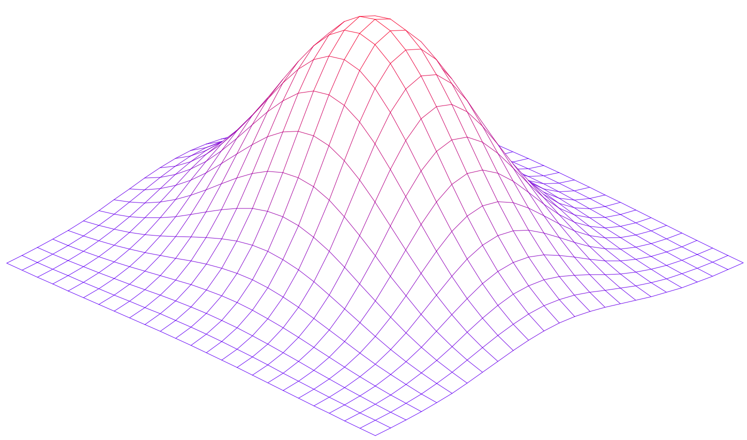
Similarly, conditional density of Y given $X = x$ is

$$f_{Y|X}(y) = \frac{f(x, y)}{f_X(x)}.$$

Interpretation: we take a section of $f(x, y)$ at x. Since

$$\int f(x, y)dy \neq 1$$

we divide by $\int f(x, y)dy$ to make up a density function (area under the density must be 1). But $\int f(x, y)dy = f_X(x)$.



A joint density, the conditional density is a section at one argument fixed.

If X and Y are independent then

$$conditional = marginal.$$

We can treat $f_{Y|X}(y)$ as usual density, i.e. we can find expectation (mean), variance, etc.

Example. A two-dimensional (bivariate) distribution is defined on the square $[0,1] \times [0,1]$ as

$$f(x,y) = \begin{cases} c(x+y-xy) \text{ if } 0 < x < 1, 0 < y < 1 \\ 0 \text{ elsewhere.} \end{cases}$$

Find appropriate constant c, marginal density $f_X(x)$ conditional density $f_{Y|X}(y)$. Are X and Y independent? What is the probability $\Pr(Y < 1/2 | X = 2/3)$?

Solution. We have

$$\int_0^1 \int_0^1 (x+y-xy)dxdy = \int_0^1 \left(\int_0^1 (x+y-xy)dx \right) dy$$

But

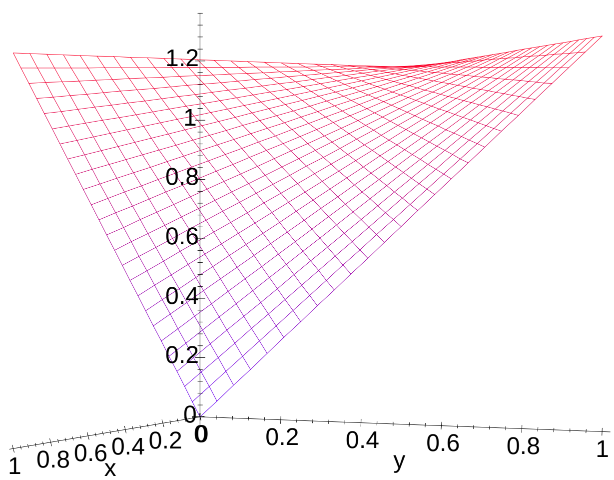
$$\int_0^1 (x+y-xy)dx = y + (1-y) \int_0^1 xdx = y + \frac{1}{2}(1-y) = \frac{y+1}{2}.$$

Then,

$$\int_0^1 \left(\int_0^1 (x+y-xy)dx \right) dy = \frac{1}{2} \int_0^1 (y+1)dy = \frac{1}{2} \times \frac{3}{2} = \frac{3}{4}.$$

Therefore, c must be $4/3$ to satisfy

$$\int \int f(x,y)dxdy = 1.$$



Joint density $\frac{4}{3}(x+y-xy)$

Marginal distribution is

$$\begin{aligned} f_X(x) &= \frac{4}{3} \int_0^1 f(x,y)dy = \frac{4}{3} \int_0^1 (x+y-xy)dy = \frac{4}{3} \left(x + (1-x) \int_0^1 ydy \right) \\ &= \frac{4}{3} \left(x + \frac{1}{2}(1-x) \right) = \frac{2}{3}(1+x) \end{aligned}$$

for $0 < x < 1$.

Conditional density is

$$f_{Y|X}(y) = \frac{f(x,y)}{f_X(x)} = 2\frac{x+y-xy}{1+x}.$$

where x is fixed and $0 < y < 1$.

To find

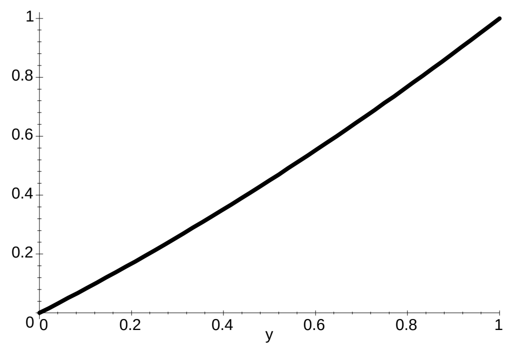
$$\Pr\left(Y < \frac{1}{2}\middle| X = \frac{2}{3}\right)$$

we need to find the conditional d. f. $F_{Y|X}(y)$ which is (x fixed)

$$F(y|x) = \int_0^y f_{Y|X}(t)dt = 2\int_0^y \frac{x+t-xt}{1+x}dt = \frac{2}{(1+x)}\int_0^y (x+t-xt)dt = y\frac{2x+y(1-x)}{1+x}.$$

If $x = 2/3$ we have

$$F(y|x=2/3) = \frac{1}{5}y(4+y)$$



Conditional d.f. for $x = 1.5$

The needed probability is

$$\Pr(Y < .5) = \frac{1}{5} \times \frac{1}{2}(4 + \frac{1}{2}) = .45$$

5. Distribution of maximum and minimum

Let $X_1, X_2, ..., X_n$ be iid. It is easy to find the distribution function of minimum, $\min X_i$ and maximum, $\max X_i$.

Indeed, let F be the common d.f.. Then, the d.f. of the maximum is

$$\begin{aligned} F_{\max}(x) &= \Pr(\max X_i \leq x) \\ &= \Pr(X_1 \leq x, X_2 \leq x, ..., X_n \leq x) \\ &= \Pr(X_1 \leq x) \cdot \Pr(X_2 \leq x) \cdots \Pr(X_n \leq x) \\ &= F^n(x). \end{aligned}$$

The density is

$$f_{\max}(x) = nF^{n-1}(x)f(x).$$

Similarly, for the minimum

$$\begin{aligned} F_{\min}(x) &= \Pr(\min X_i \leq x) \\ &= 1 - \Pr(\min X_i > x) \\ &= 1 - \Pr(X_1 > x) \cdot \Pr(X_2 > x) \cdots \Pr(X_n > x) \\ &= 1 - (1 - F(x))^n, \end{aligned}$$

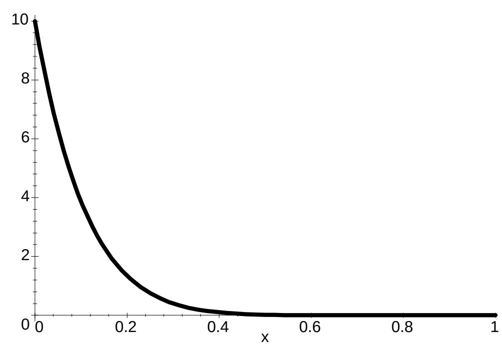
with the density

$$f_{\min}(x) = n(1 - F(x))^{n-1} f(x).$$

Problem. X_i, $i = 1, 2, ..., 10$ are iid uniformly distributed on (0,1). What is the density of $\min X_i$?

Solution. Since $F(x) = x$ we have

$$f_{\min}(x) = 10(1 - x)^9.$$



The density of the minimum, $10(1 - x)^9$

Party problem. You are invited to a party held in Keene, NH (approximately 50 miles from Hanover) starting at 7 P.M. Assume the speed of drivers on I91 follows the normal distribution with 60 ml/hour and SD 10 ml/hour, and you cannot pass. You left Hanover at 6 P.M.. and there are 5 drivers on I91 moving toward Keene. What is the probability you will be late to the party assuming you go with the maximum speed?

Solution. Let Z be the maximum speed you can go without passing other cars. This speed is the minimum of the speeds of others driver. Let X_i denote the speed of the ith driver. It is assumed $X_i \sim N(60, 10^2)$ and $X_1, ..., X_5$ are independent. Thus, your speed is

$$Z = \min_{i=1,..,5} X_i.$$

You will be late if $50/Z > 1$, i.e. if $Z < z = 50$. This probability is

$$F_Z(z) = 1 - (1 - \Phi(z; 60, 100))^5$$

where $\Phi(z; 60, 100)$ denotes the normal d.f. with mean 60 and variance 10^2. Use Table 2, A7 Rice, to find

$$\Phi(50; 60, 100) = \Phi\left(\frac{50 - 60}{10}\right) = \Phi(-1) = 1 - .8413.$$

Hence, you will be late with probability

$$1 - .8413^5 = .57854,$$

O-o.!

How many cars you should pass to be on time with probability .7?

Clearly if there are n cars ahead your probability to be on time is $.8413^n$, so that we have to find such n that

$$.8413^n = .7$$

which gives

$$n = \frac{\ln(.7)}{\ln(.8413)} = 2.064.$$

Answer: you have to pass 3 cars so that 2 cars will be ahead to be on time.

6. Expectation of multivariate RVs

Let g be any function of two arguments, its expectation (or expected value) is defined as

$$E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy.$$

The most important expected (marginal) values are mean and variance. Marginal mean is defined as

$$E(X) = \mu_X = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy = \int_{-\infty}^{\infty} xf_X(x)dx$$

and marginal variance is defined as

$$var(X) = \sigma_X^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)^2 f(x,y)dxdy = \int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)dx.$$

The same for μ_Y and σ_Y^2.

All properties of expectation remain (expectation behaves as linear function).

Theorem 6.1. *If X and Y are independent and r and s are any functions then $r(X)$ and $s(Y)$ are independent.*

Theorem 6.2. *If X and Y are independent RVs then for any functions g and h we have*

$$E(g(X)\cdot h(Y)) = E(g(X))\cdot E(h(Y)).$$

For independent variables the expectation of the product is the product of the expectations:

$$E(XY) = E(X)E(Y).$$

We prove the latter:

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

Definition 6.3. *Covariance of two RV X and Y is defined as*

$$cov(X,Y) = E(X-\mu_X)(Y-\mu_Y)$$

where μ_X and μ_Y are means of X and Y respectively.

If means are zero,

$$cov(X,Y) = E(XY).$$

Also, it is easy to see that

$$\begin{aligned} cov(X,Y) &= cov(Y,X), \\ cov(\alpha X,Y) &= \alpha \cdot cov(X,Y), \\ var(X) &= cov(X,X). \end{aligned}$$

Theorem 6.4. *Covariance of two independent variables is zero:*

$$cov(X,Y) = 0.$$

In this case X and Y are called uncorrelated.

Proof. Since X and Y are independent, $X-\mu_X$ and $Y-\mu_y$ are independent as well. Thus,

$$cov(X,Y) = E(X-\mu_x)(Y-\mu_y) = E(X-\mu_x)E(Y-\mu_y) = 0.$$

Theorem 6.5. *For any RVs X, Y, Z we have*

$$cov(X, Y+Z) = cov(X,Y) + cov(X,Z)$$

i.e. covariance behaves as a linear function.

Proof.

$$cov(X, Y+Z) = E\left((X-\mu_X)(Y+Z-\mu_{Y+Z})\right)$$

But

$$\mu_{Y+Z} = E(Y+Z) = E(Y) + E(Z) = \mu_Y + \mu_Z.$$

Thus,

$$\begin{aligned} & E\left((X-\mu_X)(Y+Z-\mu_{Y+Z})\right) \\ = \; & E\left((X-\mu_X)((Y-\mu_Y)+(Z-\mu_Z))\right) \\ = \; & E\left((X-\mu_X)(Y-\mu_Y)\right) + E\left((X-\mu_X)(Z-\mu_Z)\right) \\ = \; & cov(X,Y) + cov(X,Z). \end{aligned}$$

Problem. Find the variance of XY where X and Y are independent and $E(X) = E(Y) = 0$ and $var(X) = var(Y) = \sigma^2$.

Solution. Let us denote $Z = XY$. We have

$$E(Z) = E(XY) = E(X)E(Y) = 0 \cdot 0 = 0.$$

Find the second (noncentral moment)

$$E(Z^2) = E(XY)^2 = E(X^2Y^2) = E(X^2)E(Y^2)$$

But

$$E(X^2) = var(X) + (E(X))^2 = \sigma^2 + 0^2 = \sigma^2.$$

Therefore,

$$E(Z^2) = \sigma^2 \cdot \sigma^2 = \sigma^4,$$

and

$$var(Z^2) = E(Z^2) - (E(Z))^2 = \sigma^4 - 0^2 = \sigma^4.$$

For any X and Y

$$var(X+Y) = var(X) + 2cov(X,Y) + var(Y).$$

Proof. Let $Z = X + Y$. Then

$$\mu_Z = E(Z) = E(X+Y) = E(X) + E(Y) = \mu_X + \mu_Y.$$

Then,

$$\begin{aligned}
var(X+Y) &= var(Z) = E(Z-\mu_Z)^2 = E\left[(X+Y) - (\mu_X+\mu_Y)\right]^2 \\
&= E\left[(X-\mu_X) + (Y-\mu_Y)\right]^2 \\
&= E\left[(X-\mu_X)^2 + 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2\right] \\
&= E(X-\mu_X)^2 + 2E(X-\mu_X)(Y-\mu_Y) + E(Y-\mu_Y)^2 \\
&= var(X) + 2cov(X,Y) + var(Y).
\end{aligned}$$

Fact to remember: If X and Y are uncorrelated (or independent) then

$$var(X+Y) = var(X) + var(Y).$$

7. Application to finance – optimal investment

Two independent stocks have the same expected performance (return) but different risk. To diversify, what should be the proportion of each stock in your portfolio to minimize the overall risk.

We denote X – the first stock, and Y – the second stock. Thus, it is assumed that stock return is a RV. Expected performance is expected value,

$$E(X) = E(Y) = \mu_X = \mu_Y = \mu.$$

Then, the variance is associated with risk. Denote

$$var(X) = \sigma_X^2, \quad var(Y) = \sigma_Y^2$$

The independence implies

$$cov(X,Y) = 0.$$

We diversify, which is formalized as follows

$$Z = \alpha X + (1-\alpha)Y$$

a new stock where α is the proportion to find, $0 < \alpha < 1$.

Question #1. Does Z have the same expected return? Yes:

$$\begin{aligned} E(Z) &= E(\alpha X + (1-\alpha)Y) = E(\alpha X) + E((1-\alpha)Y) \\ &= \alpha E(X) + (1-a)E(Y) = \alpha\mu + (1-\alpha)\mu = \mu. \end{aligned}$$

What about Z's risk?

$$\begin{aligned} var(Z) &= var(\alpha X + (1-\alpha)Y) = var(\alpha X) + var((1-\alpha)Y) \\ &= \alpha^2 var(X) + (1-\alpha)^2 var(Y) = \alpha^2\sigma_X^2 + (1-\alpha)^2\sigma_Y^2. \end{aligned}$$

We need to minimize the risk, i.e.

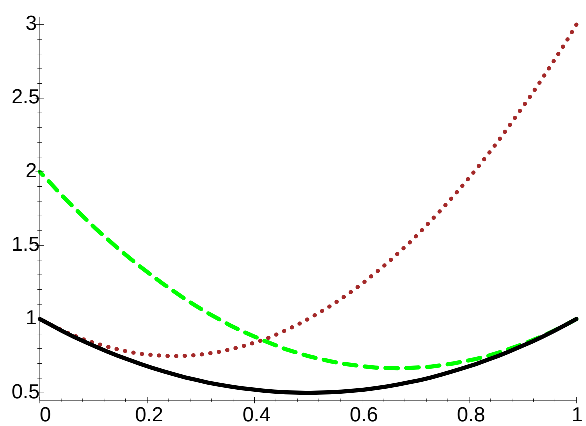
$$var(Z) \Rightarrow \min_\alpha$$

i.e.,

$$\alpha^2\sigma_X^2 + (1-\alpha)^2\sigma_Y^2 \Rightarrow \min_\alpha$$

Denote

$$R(\alpha) = \alpha^2\sigma_X^2 + (1-\alpha)^2\sigma_Y^2$$



Function $R(\alpha)$ for different stock variances.
Solid: $\sigma_X^2 = \sigma_Y^2 = 1$,
Dash: $\sigma_X^2 = 1, \sigma_Y^2 = 2$,
Dot: $\sigma_X^2 = 3, \sigma_Y^2 = 1$

Find the minimum of $R(\alpha)$:

$$\frac{dR}{d\alpha} = 2\alpha\sigma_X^2 - 2(1-\alpha)\sigma_Y^2 = 0$$

that gives

$$\alpha_{opt} = \frac{\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}. \tag{7.1}$$

Minimum risk is provided by (7.1), and is equal to

$$
\begin{aligned}
& R_{\min} \\
= \; & \frac{\sigma_Y^4}{(\sigma_X^2+\sigma_Y^2)^2}\sigma_X^2 + \frac{\sigma_X^4}{(\sigma_X^2+\sigma_Y^2)^2}\sigma_Y^2 \\
= \; & \frac{\sigma_X^2\sigma_Y^2}{(\sigma_X^2+\sigma_Y^2)^2}\left(\sigma_X^2+\sigma_Y^2\right) = \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2+\sigma_Y^2}
\end{aligned}
$$

Show that

$$R_{\min} < R_X \text{ and } R_{\min} < R_Y.$$

The best proportion is

$$\frac{\alpha_{opt}}{1-\alpha_{opt}} = \frac{1/\sigma_X^2}{1/\sigma_Y^2}.$$

Rules:
1. Always diversify.
2. If stocks have the same performance, independent, and equally risky use the 50/50 rule.
3. The optimal stock proportion is reciprocal of variances.

8. Homework (due February 3)

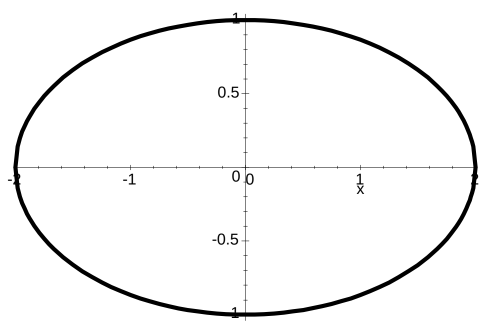
Maximum number of point.

1. (5 points). Solve 3.8.6 from Rice (p. 104).

Solution. The joint density of (X, Y) is constant (c) within the ellips

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$$

and zero outside.



Ellips for $a = 2$ and $b = 1$

The constant c must satisfy

$$1 = \iint\limits_{x^2/a^2+y^2/b^2\leq 1} c\,dxdy = c \iint\limits_{x^2/a^2+y^2/b^2\leq 1} dxdy = c4\frac{b}{a}\int_0^a \sqrt{a^2-x^2}dx.$$

To find the integral we change the variable: $x = a\sin t$. The upper limit of t is $\pi/2$ and the lower limit is 0. Then we notice that $dx = a\cos t dt$ and $\sqrt{a^2 - x^2} = a\sqrt{1-\sin^2 t} = a\cos t$. Hence,

$$\int_0^a \sqrt{a^2-x^2}dx = a^2\int_0^{\pi/2}\cos^2 t dt = \frac{a^2}{2}\int_0^{\pi/2}(1+\cos 2t)dt = \frac{a^2}{2}\frac{\pi}{2}.$$

Thus, the area of the ellips is πab and we obtain $c = 1/(\pi ab)$.

First, we find the marginal density for X. When $X = x$ is fixed them y runs from $-\frac{b}{a}\sqrt{a^2-x^2}$ to $\frac{b}{a}\sqrt{a^2-x^2}$. Hence, to find the marginal density we integrate out y as follows

$$f_X(x) = \frac{1}{\pi ab}\int_{-\frac{b}{a}\sqrt{a^2-x^2}}^{\frac{b}{a}\sqrt{a^2-x^2}} dy = \frac{2\frac{b}{a}\sqrt{a^2-x^2}}{\pi ab} = \frac{2\sqrt{a^2-x^2}}{\pi a^2}$$

for $-a < x < a$ and zero outside $(-a, a)$. Similarly,

$$f_Y(y) = \frac{1}{\pi ab}\int_{-\frac{a}{b}\sqrt{a^2-y^2}}^{\frac{a}{b}\sqrt{b^2-y^2}} dx = \frac{2\frac{a}{b}\sqrt{b^2-y^2}}{\pi ab} = \frac{2\sqrt{b^2-y^2}}{\pi b^2}, \quad -b < y < b.$$

2. (6 points). Solve 3.8.19 from Rice (p.105).

Solution. RV T_1 has the density function $\alpha e^{-\alpha x}$ and T_2 has the density $\beta e^{-\beta y}$. The joint density is the product, $f(x,y) = \alpha b e^{-\alpha x}e^{-\beta y}$. The first probability is

$$\Pr(T_1 > T_2) = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{x} f(x,y)dy\right)dx.$$

Since exponential distribution is defined for positive values we obtain

$$\begin{aligned}
\Pr(T_1 &> T_2) = \int_0^\infty \alpha\left(\beta\int_0^x e^{-\alpha x}e^{-\beta y}dy\right)dx = \int_0^\infty \alpha e^{-\alpha x}\left(\beta\int_0^x e^{-\beta y}dy\right)dx \\
&= \int_0^\infty \alpha e^{-\alpha x}\left(1-e^{-\beta x}\right)dx = \int_0^\infty \alpha e^{-\alpha x}dx - \alpha\int_0^\infty e^{-(\alpha+\beta)x}dx = 1 - \frac{\alpha}{\alpha+\beta} = \frac{\beta}{\alpha+\beta}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\Pr(T_1 &> 2T_2) = \int_0^\infty \alpha\left(\beta\int_0^{x/2} e^{-\alpha x}e^{-\beta y}dy\right)dx = \int_0^\infty \alpha e^{-\alpha x}\left(\beta\int_0^{x/2} e^{-\beta y}dy\right)dx \\
&= \int_0^\infty \alpha e^{-\alpha x}\left(1-e^{-\beta x/2}\right)dx = \int_0^\infty \alpha e^{-\alpha x}dx - \alpha\int_0^\infty e^{-(\alpha+\beta/2)x}dx = 1 - \frac{\alpha}{\alpha+\beta/2} = \frac{\beta}{2\alpha+\beta}.
\end{aligned}$$

3. (7 points). If X is the amount of money (in dollars) that a salesperson spends on gasoline during a day, and Y is the corresponding amount of money for which he/she is reimbursed, the joint density of these two random variables is given by

$$f(x,y) = \begin{cases} \frac{1}{25}\left(\frac{20-x}{x}\right) & for \qquad 10 < x < 20, .5x < y < x \\ 0 & elsewhere \end{cases}$$

find: (a) the marginal density of X, (b) the conditional density of Y given $X = 12$, (c) the probability that the salesperson will be reimbursed at least \$8 when spending \$12.

Solution. To calculate marginal density we fix x and then y belongs to the interval $(x/2, x)$. Thus, the marginal density is

$$\begin{aligned} f_X(x) &= \int_{x/2}^{x} f(x,y)dy = \int_{x/2}^{x} \frac{1}{25}\left(\frac{20-x}{x}\right)dy = \frac{1}{25}\left(\frac{20-x}{x}\right)\int_{x/2}^{x} dy \\ &= \frac{1}{25}\left(\frac{20-x}{x}\right)\frac{x}{2} = \frac{20-x}{50} \text{ for } 10 < x < 20. \end{aligned}$$

The conditional density is $f(x,y)/f_X(x)$ which leads to a uniform distribution with the density

$$f_{Y|X}(y) = \frac{2}{x} \text{ for } .5x < y < x.$$

If $X = 12$ then the conditional density becomes $1/6$ for $6 < y < 12$. The needed probability is $\Pr(Y \geq 8|X = 12)$. Since $f_{Y|X}(y) = 1/6$ this probability is (12-8)/6=4/6=2/3.

4. (4 points). Let X and Y are two independent RVs with means and variance μ_X, σ_X^2 and μ_Y, σ_Y^2 respectively. Find the variance of the product, XY.

Solution. By th definition, $var(XY) = E(XY)^2 - E^2(XY)$. But due to independence

$$E(XY)^2 = E(X^2Y^2) = E(X^2)E(Y^2) = (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2).$$

But again due to independence $E(XY) = \mu_X\mu_Y$ so that we obtain

$$var(XY) = (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2 = \sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2.$$

5. (5 points). You are taking part in the auction and there are 10 other people. Each person is willing to pay from \$1,000 to \$2,000 with equal preference (uniform distribution). You can spend maximum \$1,990. What is the probability to get the object of the auction?

Solution. For each participant $X_i \sim U(1000, 2000)$ with the distribution function $F(x) = 1000^{-1}(x - 1000) = (x/1000 - 1)$. You get the object of the auction if $\max X_i < 1990$. But the distribution of the maximum is

$$F^{10}(x) = (x/1000 - 1)^{10}.$$

Substituting $x = 1990$ we obtain the asked probability,

$$(1990/1000 - 1)^{10} = .9$$

6. (3 points). Prove that the minimum of iid exponentially distributed RVs is again an exponentially distributed RV. Is this true for the maximum?

Solution. The distribution function of each X_i is $1-\exp(\lambda x)$. The distribution function of $\min X_i$ is

$$1 - (1 - F(x))^n = 1 - \exp(-\lambda n x),$$

which is again an exponential distribution with the rate λn. It is not true for the maximum because $(1 - \exp(-\lambda x))^n$ is not of the form $1 - \exp(-\alpha x)$.

7. (6 points). How to minimize risk in optimal investment problem if the stock returns are not the same?

Solution. The optimal rule remains the same as for equal returns because $var(\alpha X + (1-\alpha)Y) = \alpha^2 var(X) + (1-\alpha)^2 var(Y)$ which does not contain means of X and Y.

8. (3 points). There are three independent RVs X, Y, Z. One generates a couple of new RVs as follows, $U = X + Y$ and $V = Y + Z$. Are U and V independent? Show the work.

Solution. No, U and V are dependent. To prove this we calculate the covariance,

$$\begin{aligned} cov(U,V) &= cov(X+Y,Y+Z) = cov(X,Y) + cov(X,Z) + cov(Y,Y) + cov(Y,Z) \\ &= 0 + 0 + var(Y) + 0 = var(Y) > 0 \end{aligned}$$

Therefore, $cov > 0$ and U and V cannot be independent.

9. (3 points). RVs X and Y are independent, are $X - Y$ and $X + Y$ independent? Show the work.

Solution. Again, we calculate covariance

$$cov(X-Y, X+Y) = cov(X,X) + cov(X,Y) - cov(Y,X) - cov(Y,Y) = var(X) - var(Y).$$

Generally, $X - Y$ and $X + Y$ are dependent. They are uncorrelated only if $var(X) = var(Y)$.

10. (4 points). Prove that if two RVs are uncorrelated then linear functions of them are uncorrelated as well (if X and Y are uncorrelated then $aX + b$ and $cY + d$ are uncorrelated).

Solution. Calculate the covariance

$$cov(aX + b, cY + d) = cov(aX, cY) + cov(aX, d) + cov(b, cY) + cov(b, d).$$

But covariance between RV and constant is zero. It implies that

$$cov(aX + b, cY + d) = cov(aX, cY) = ac \times cov(X, Y) = 0$$

11. (5 points). X_i are iid with variance σ^2. Find the variance of the average, $\sum_{i=1}^n X_i/n$.

Solution. Since X_i are iid

$$\begin{aligned} var(\frac{1}{n}\sum_{i=1}^n X_i) &= \frac{1}{n^2}\sum_{i=1}^n var(X_i) = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 \\ &= \frac{\sigma^2}{n}. \end{aligned}$$