HG
03.09.14

Errata to lecture 5 (Wednesday 3 Sept.)

There were some mess-ups in the lecture – corrected below

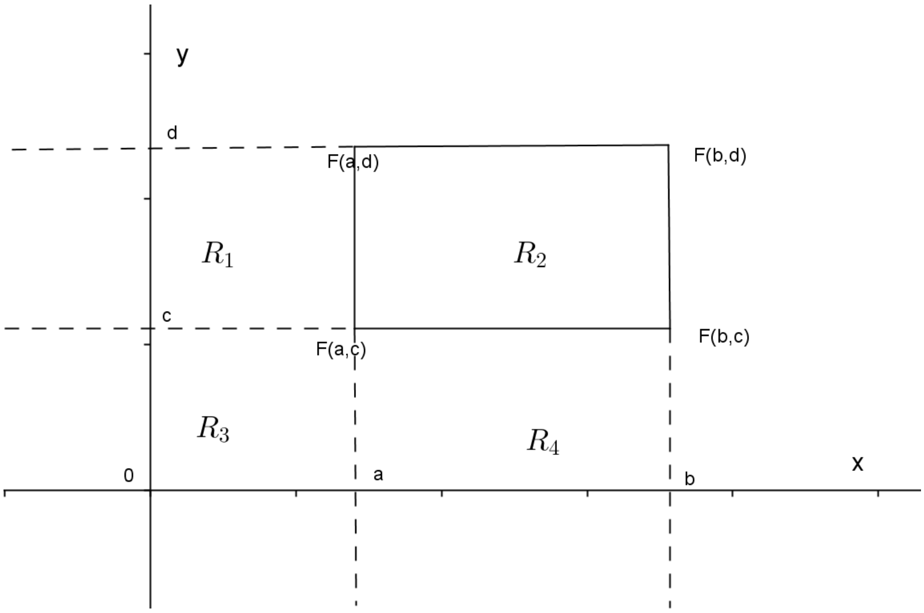
1. **Correction to the probability of a rectangle based on the joint cdf *F*.**
   (I mixed up some of the arguments in the lecture, so correct your notes)

Let the rv pair (X,Y) have the cdf $F(x,y) = P\left[(X \le x) \cap (Y \le y)\right]$

Then, for the rectangle R_2 in the figure below, we have

$$P((X,Y) \in R_2) = P\left[(a < X \le b) \cap (c < Y \le d)\right] = F(b,d) - F(b,c) - F(a,d) + F(a,c)$$

Proof



Write the volume under the pdf over any rectangle R as $v(R) = \mathrm{P}\left[(X,Y) \in R\right]$, so the probability we want is $v(R_2)$.

Now the total rectangle from $-\infty$ up to the point (b,d) is composed by 4 disjoint parts, R_1, R_2, R_3, R_4. Hence $F(b,d) = v(R_1) + v(R_2) + v(R_3) + v(R_4)$.

Similarly, $F(b,c) = v(R_3) + v(R_4)$, which gives

$F(b,d) - F(b,c) = v(R_1) + v(R_2)$, i.e., the volume over the upper strip in the figure. From the figure we also see that $v(R_1) = F(a,d) - F(a,c)$, and, hence

$v(R_2) = F(b,d) - F(b,c) - \left(F(a,d) - F(a,c)\right) = F(b,d) - F(b,c) - F(a,d) + F(a,c)$ **(End of proof)**

2. Let (X,Y) be a pair of continuous rv's with cdf $F(x,y) = P\left[\left(X \le x\right) \cap (Y \le y)\right]$
In the proof of the fact that the cdf, *F*, determines the pdf, *f*, by double derivation, I came out wrong in the lecture, so here a correct version follows:

The important result to be proven is: $f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y) = \frac{\partial^2}{\partial y \partial x} F(x,y) \overset{\text{def}}{=} \frac{\partial}{\partial y}\left[\frac{\partial}{\partial x} F(x,y)\right]$

(Remember from math courses that the order of differentiation does not matter, i.e., we may first differentiate with respect to *x* and then with respect to *y* (as here), or opposite.)

Proof.
The result follows from the following rule that we have used before:

(*) $\boxed{\frac{\partial}{\partial x}\int_a^x g(u)du = g(x) \text{ for any integrable } g(u)}$

We have $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)dvdu = \int_{-\infty}^{x}\left[\int_{-\infty}^{y} f(u,v)dv\right]du = \int_{-\infty}^{x} J(u,y)du$, where $J(u,y)$ is the inner integral, $J(u,y) = \int_{-\infty}^{y} f(u,v)dv$. (Remember that *u* and *y* are fixed (freely chosen) values when doing the inner integration with respect to *v*. The result of the inner integration must therefore depend on *u* and *y*, i.e., the result must be a function of *u* and *y*.)

Using (*), we get

$$\frac{\partial}{\partial x} F(x,y) = \frac{\partial}{\partial x}\int_{-\infty}^{x} J(u,y)du \overset{(*)}{=} J(x,y) = \int_{-\infty}^{y} f(x,v)dv$$

Hence, using (*) again, we get

$$\frac{\partial^2}{\partial y \partial x} F(x,y) = \frac{\partial}{\partial y}\left[\frac{\partial}{\partial x} F(x,y)\right] = \frac{\partial}{\partial y}\int_{-\infty}^{y} f(x,v)dv \overset{(*)}{=} f(x,y)$$

which is the pdf. (**End of proof**)

Note. It often occurs in practice that the cdf, *F*, is not differentiable in isolated points or lines. This does not matter since we may define the pdf, $f(x,y)$, as we like in isolated points or on lines (or curves) without affecting the distribution. The reason for this is that the distribution only comprises probabilities about the rv pair, (X,Y), and all probabilities are volumes between the (*x*,*y*)-plane and $f(x,y)$, and volumes over isolated points or lines are always zero (for continuous rv's, *X*, *Y*).