Chan and Khandani

Jointly Distributed Random Variables

Introduction

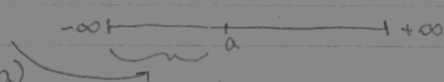
Today in ECE 316, we stuff multiple variables together and call it a new probability.

Joint Distribution Functions

A clearer name for these would probably be joint cumulative distribution functions. Remember that our original CDFs looked like this:

$$F_X(a) = P\{X \leq a\}$$

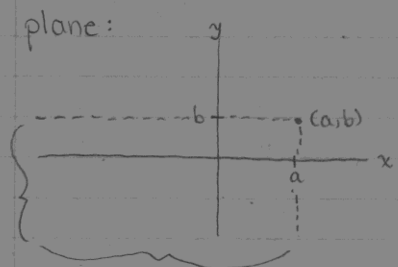
And graphically, represents the points where $x \in (-\infty, a)$



So let's introduce the joint distribution

$$F_{XY}(a, b) = P\{X \leq a, Y \leq b\}$$

We would then have to represent this in a 2D plane:



It now consists of all points from $(-\infty, a)$ for x and $(-\infty, b)$ for y.

Note that this is not describing the probability itself, rather the notable points we need to watch out for when calculating the probability itself.

Since the joint distribution consists of two cumulative distribution functions, all it requires is a clever trick to isolate each.

$$F_X(a) = P\{X \leq a\} = P\{X \leq a, Y < \infty\} = F_{XY}(a, \infty)$$

$$F_Y(b) = P\{Y \leq b\} = P\{X \leq \infty, Y \leq b\} = F_{XY}(\infty, b)$$

By passing ∞ as the parameter for the other variable, we can force that probability to become 1, essentially ignoring it.

As for the discrete analogue:

$$p(x, y) = P\{X = x, Y = y\}$$

The trick to separate the discrete version is to sum the probabilities over the entire set a variable is defined.

$$p_X(x) = P\{X = x\} = \sum_y P\{X = x, Y = y\} = \sum_y p(x, y)$$

$$p_Y(y) = P\{Y = y\} = \sum_x P\{X = x, Y = y\} = \sum_x p(x, y)$$

Let's go over a quick example to make sure we get what shenanigans we're trying to pull.

ex. We have 2 players, each with their own die. Each of them roll, and the following results are possible.

1) If both players roll an even number, they both lose 1 dollar.
2) If both players roll an odd number, they both lose 1 dollar.
3) If one player rolls an odd, and the other rolls an even, the player who rolled the odd gets 3 dollars. The player who rolled the even loses 1.

We can now define (starting at 0 dollars)

$$P_1, P_2 = \begin{cases} -1 & \text{if same roll} \\ -1 & \text{if different roll, we roll even} \\ +3 & \text{if different roll, we roll odd} \end{cases}$$

Then, we define

$$F_{P_1P_2}(a, b) = P\{P_1 \leq a,\ P_2 \leq b\},$$

the joint distribution after 10 rolls have been made. Describe what

$$F_{P_1P_2}(2, 0) = P\{P_1 \leq 2,\ P_2 \leq 0\}$$

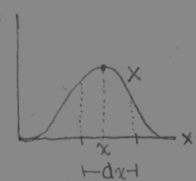
represents. What are the minimum and maximum values of P_1 / P_2?

The description is as such: what is the probability that player 1 has 2 or fewer dollars and player 2 has 0 or fewer dollars after 10 rolls?

The minimum/maximum occur from the best and worst case scenarios - 10 losses → -10 dollars, 10 wins → +30 dollars.

Jointly Continuous Random Variables

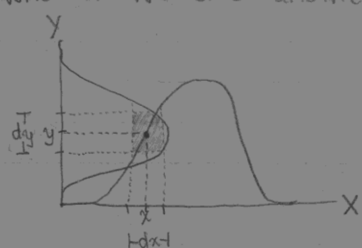
Before we get into what jointly continuous means, let's remind ourselves how the probability density of variables are defined.



Let's say we want to investigate the probability that X is within a certain range of an arbitrary constant x. We've defined the range as $\pm dx/2$ around x, so our probability density would look like this

$$f_X(x)\,dx = P\{x - \tfrac{dx}{2} \le X \le x + \tfrac{dx}{2}\}$$

So what if we add another dimension?



The Y behaves the exact same way, so looking for $y \pm dy/2$ yields:

$$f_Y(y)\,dy = P\{y - dy/2 \le Y \le y + dy/2\}$$

Merging these two creates a region of area $dxdy$. Now we can ask the probability that a point (x,y) exists in this region. Putting this together:

$$f_{XY}(x,y)\,dx\,dy = P\{x - \tfrac{dx}{2} \le X \le x + \tfrac{dx}{2},\ y - \tfrac{dy}{2} \le Y \le y + \tfrac{dy}{2}\}$$

So in order to actually calculate this, we must integrate over our specified region of $(x - dx/2,\ x + dx/2)$ and $(y - dy/2,\ y + dy/2)$.

This brings us back to our original question - what is jointly continuous? Mathematically, it's given as

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

Let's break it down.

① $P\{(X,Y)\in C\}$

This is the probability that X, Y results in values such that the point (X, Y) lies in some defined region C.

② $\iint_{(x,y)\in C}$

We're integrating over the entire area C. So if C was defined as $x\in(2,3)$ and $y\in(3,4)$, this would become

$$\int_2^3\int_3^4 \qquad \text{or} \qquad \int_3^4\int_2^3$$

depending on what order you want to do the integration.

③ $f(x,y)\,dx\,dy$

As before, this is the combined probability density of X and Y.

So the probability that (X, Y) is in the region C is found by integrating over that same region and summing the probabilities of the random variables.

However, "jointly continuous" is a POSSIBLE property OF two or more random variables.

So, as such, there are constraints on when we can actually call two random variables jointly continuous, which is largely to do with whether the integration is actually possible.

But if we think a little more closely about it, the only time an integration isn't possible is when we have undefined values - for example,

$$\int_{-5}^{10} 5/0$$

As probabilities are a reflection of real-life situations, it's impossible to have an undefined probability. So for the purpose of this course, we can just assume this property holds for all cases.

Now that we know this, we can return to the joint cumulative distribution and have some solid math instead of vague inequalities.

$$F_{XY}(a,b) = \int_{-\infty}^{b}\int_{-\infty}^{a} f(x,y)\,dx\,dy$$

<u>Independent Random Variables</u>

This is basically exactly the same as before - if the product of the individual probabilities are the same as the actual probability, they're independent:

$$P\{X \in A, Y \in B\} = P\{X \in A\}\{Y \in B\}$$

$$\therefore$$

$$F(a,b) = F_X(a)\,F_Y(b)$$

$$\therefore$$

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

Nothing fantastically exciting nor surprising here

Sums of Independent Random Variables

Sometimes we'll encounter questions that ask the sum of random variables - perhaps the sum of dice rolls, or the total accrued money of a troupe of gamblers. Either way, these random variables must be independent and continuous. The cumulative distribution of the sum is:

$$F_{X+Y}(a) = P\{X + Y \le a\}$$

$$= \iint_{(x+y)\le a} f_X(x)\, f_Y(y)\, dx\, dy$$

independent means their joint density is the product

$$= \iint_{x \le a-y} f_X(x) f_Y(y)\, dx\, dy$$

So now, we have two integrals but only one has bounds on it: x.

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

Remember that $\int_{-\infty}^{a-y} f_X(x)\, dx$ is the cumulative distribution of x itself, so we can replace that in our equation:

$$= \int_{-\infty}^{\infty} F_X(a-y)\, f_Y(y)\, dy$$

Alternatively, we could have replaced $f_Y(y)$:

$$= \int_{-\infty}^{\infty} F_Y(a-x)\, f_X(x)\, dx$$

Finally, we can find the probability density by differentiating the cumulative distribution.

$$f_{X+Y}(a) = \frac{d}{da}\int_{-\infty}^{\infty} F_X(a-y)\, f_Y(y)\, dy$$

$$= \frac{d}{da}\int_{-\infty}^{\infty}\int_{-\infty}^{a-y} f_X(a-y)\, f_Y(y)\, dx\, dy$$

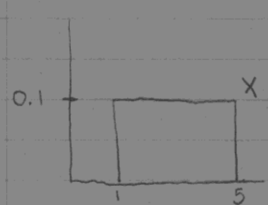
The derivative cancels out the integral that uses a:

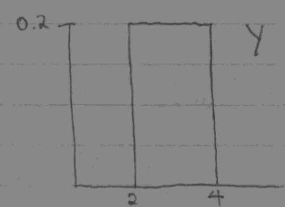
$$= \int_{-\infty}^{\infty} f_X(a-y)\, f_Y(y)\, dy$$

or

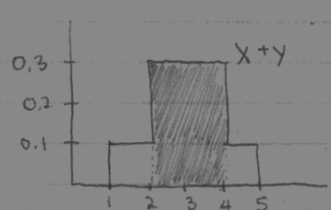
$$= \int_{-\infty}^{\infty} f_Y(a-x)\, f_X(x)\, dx$$

We can take a quick look to see why this works.





We've got two distributions here (ignore that the totals aren't actually 1, I can't be arsed to do that, so pretend)



Let's say $a = 0.3$.

$\int_{-\infty}^{\infty} f_X(x)\, dx$ is just 1.

$$\int_{-\infty}^{\infty} f_Y(0.3-x)\, dx = \int_{-\infty}^{\infty} f_Y(0.3-0.1)\, dx$$

0.1 is the only value of X.

$\therefore \int_{-\infty}^{\infty} f_Y(0.2)\, dx$ is the entirety of Y: from 2 to 4.

Sum of Gamma Random Variables

There's a special case for independent random gamma variables because the math is pretty spooky. They must share a variance λ: that is, their parameters are (r_1, λ) and (r_2, λ).

$X+Y$ becomes a random gamma variable with the parameters (r_1+r_2, λ). I'm going to skip the proof because it's too difficult for me to understand.

$$\therefore f_{X+Y}(a) = \frac{\lambda e^{-\lambda a} (\lambda a)^{r_1+r_2-1}}{\Gamma(r_1+r_2)}$$

where $\Gamma(x)$ is the gamma function $\int_0^\infty t^{x-1} e^{-t}\,dt$.

Sum of Square of Standard Normal Random Variables

If we've got a bunch of standard normals Z_1 to Z_n,

$$Y \equiv \sum_{i=1}^{n} (Z_i)^2$$

has a "chi-squared" (χ^2) distribution, with n degrees of freedom. If $n=1$,

$$f_{Z^2}(y) = \frac{(1/2) e^{-(1/2)y} ((1/2)y)^{1/2-1}}{\Gamma(1/2)}$$

Which is a gamma distribution with parameters $(1/2, 1/2)$. We can then draw the conclusion that a chi squared distribution with n degrees of freedom is given by

$$f_{\chi^2}(y) = \frac{e^{-y/2}\, y^{n/2-1}}{2^{n/2}\,\Gamma(n/2)},$$

a gamma distribution with parameters $(n/2, 1/2)$.

$$\Gamma(n/2) \xrightarrow{\text{even}} \left[\left(\tfrac{n}{2}\right)-1\right]!$$

$\xrightarrow{\text{odd}}$ use $\Gamma(t) = (t-1)\Gamma(t-1)$ and $\Gamma(1/2) = \sqrt{\pi}$

Sum of Normal Random Variables

Adding things with a lot of terms can get very complex, so knowing that the sum of random normal variables is still a normal random variable can save us a lot of grief.

$\sum_{i=1}^{n} X_i$ is normally distributed with parameters

$$\sum_{i=1}^{n} \mu_i \quad \text{and} \quad \sum_{i=1}^{n} \sigma^2$$

So you can just add up all the means and all of the variances and call it a day.

Conditional Distributions (Discrete)

Originally given as

$$P(E \mid F) = \frac{P(EF)}{P(F)},$$

we need to then express it as a function of the discrete probability mass in order to find the joint probability

$$p_{X|Y}(x|y) = P\{X = x \mid Y = y\}$$

Here, we've done just that: the probability that $X = x$ given that $Y = y$.

$$= \frac{P\{X = x,\ Y = y\}}{P\{Y = y\}}$$

And here we've expanded it to take the same form as the original conditional probability.

$$= \frac{p(x, y)}{p_Y(y)}$$

Now, we can collapse the numerator into the joint probability, and the denominator into a singular probability.

Then, if we wanted the cumulative distribution instead, we ask "What is the probability that $X \leq x$ given that $Y = y$?", therefore

$$F_{X|Y}(x|y) = P\{X \leq x \mid Y = y\}$$

$$= \sum_{a \leq x} p_{X|Y}(a|y)$$

↑ sum of the individual probabilities

$$= P\{X = x\}$$

↑ if independent

Conditional Distribution: (Continuous)

By the same virtue as the discrete case,

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

Joint Probability Distribution Functions

So let's say we have two random variables X_1 and X_2, with the joint density f_{X_1,X_2}.

If we have two other variables which are functions of X_1 and X_2, say:

$$Y_1 = g_1(X_1, X_2) \qquad Y_2 = g_2(X_1, X_2)$$

How do we find the joint distribution of Y_1 and Y_2: f_{Y_1,Y_2}? It's given by a specific equation if some conditions are true

1) You can solve Y_1 and Y_2 for X_1 and X_2 by some relation ship

$$x_1 = h_1(y_1, y_2) \qquad x_2 = h_2(y_1, y_2).$$

Let's do a quick example.

$$Y_1 = 3X_1 + 4X_2 \qquad Y_2 = 6X_1 - X_2$$

Let's solve the first equation for X_1.

$$Y_1 = 3X_1 + 4X_2$$
$$\frac{Y_1 - 4X_2}{3} = X_1$$

Plug this into the second equation:

$$Y_2 = 6\left(\frac{Y_1 - 4X_2}{3}\right) - X_2$$
$$Y_2 = \frac{6Y_1 - 24X_2 - 3X_2}{3}$$
$$Y_2 = 2Y_1 - \frac{27X_2}{3}$$
$$Y_2 - 2Y_1 = \frac{-27X_2}{3}$$
$$X_2 = \frac{3Y_2 - 6Y_1}{-27}$$

We now have an equation for X_2 → now we need one for X_1.

$$X_1 = \frac{Y_1 - 4X_2}{3}$$
$$= \frac{Y_1 - 4\left[\frac{3Y_2 - 6Y_1}{-27}\right]}{3}$$
$$= \left(Y_1 - \frac{12Y_2 + 24Y_1}{-27}\right) / 3$$
$$= \left(\frac{-27Y_1 - 12Y_2 + 24Y_1}{-27}\right) / 3$$
$$= \frac{Y_1 - 4Y_2}{27}$$

So for the Y_1 and Y_2 we've defined, this condition is true.

2) The functions g_1 and g_2's partial derivatives must exist for all possible (x_1, x_2) (so you can't say $Y = x_1/|x_2|$ if 0 is a valid x_2, as it isn't differentiable there), AND the Jacobian $\neq 0$.

Remember that the Jacobian is a determinant of partial derivatives:

$$J(x_1, x_2) \equiv \frac{\partial g_1}{\partial x_1}\frac{\partial g_2}{\partial x_2} - \frac{\partial g_1}{\partial x_2}\frac{\partial g_2}{\partial x_1} \neq 0.$$

Let's try this out with the functions from before.

$$\cancel{\frac{\partial g_1}{\partial x_1}(3x_1 + 4x_2) = 3} \qquad \cancel{\frac{\partial g_2}{\partial x_2}(6x_1 - x_2) = -1}$$ *

$$\cancel{\frac{\partial g_1}{\partial x_2}(3x_1 + 4x_2) = 4} \qquad \cancel{\frac{\partial g_2}{\partial x_1}(6x_1 - x_2) = 6}$$

$$\cancel{\therefore J(x_1, x_2) = (3)(-1) - (4)(6)}$$
$$\cancel{= -3 - 24}$$
$$\cancel{= -27}$$

So this condition is true. If both conditions are true, the joint distribution of Y_1 and Y_2 is given as

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2)\,|J(x_1, x_2)|^{-1} \leftarrow \text{inverse}$$

where $x_1 = h_1(y_1, y_2)$
$x_2 = h_2(y_1, y_2)$

For us, this final distribution would look like:

* this example is wrong - see example 21 instead

*see example 21 instead

$$\cancel{f_{Y_1,Y_2}(y_1,y_2) = \frac{f_{X_1,X_2}\left(\frac{y_1-4y_2}{-27}, \frac{3y_2-6y_1}{-27}\right)}{27}}$$

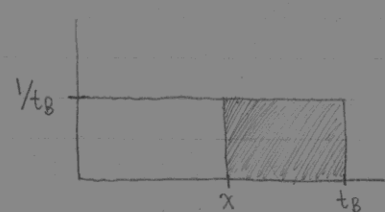
Finally, done with the material. Example time!

ex 2) We get to a bus stop at $t=0$. There are two buses running, where their arrival times A and B are uniformly distributed on $[0, t_a]$ and $[0, t_b]$, respectively, where $t_A \leq t_B$.

Their arrivals are independent, much like how real buses show up whenever the fuck they want. Find the probability that the A bus arrives first

What we're trying to find here is the probability that A shows up, where B has not shown up, so this is actually a conditional probability

$\therefore$ If A arrives at some time x, (which is $\leq t_a$) it is first iff B arrives between x and t_B



So the probability of B arriving between x and t_B is $(t_B - x)(1/t_B)$, the area of that portion.

Even though the condition is present in this probability, the events are independent, so we can say

$$P\{B \text{ comes after } x \mid A \text{ arrives at } 0 \leq x \leq t_A\}$$

is simply

$$= \frac{t_B - x}{t_B}$$
$$= 1 - \frac{x}{t_B}$$

Now, this is for just one specific $x \rightarrow$ we want to look at all possible x; that means everything from 0 to t_A, all possible arrival times of A.

$$P\{A < B\} = \int_{-\infty}^{\infty} \left(1 - \frac{x}{t_B}\right)\left(\frac{1}{t_A}\right) dx$$

probability of arriving at a specified x

$$= \int_0^{t_A} \frac{1}{t_A} - \frac{x}{t_A t_B}\, dx$$
$$= \int_0^{t_A} \frac{1}{t_A} dx - \int_0^{t_A} \frac{x}{t_A t_B} dx$$
$$= \left[\frac{x}{t_A}\right]_0^{t_A} - \left[\frac{x^2}{2t_A t_B}\right]_0^{t_A}$$
$$= [1 - 0] - \left[\frac{t_A}{2t_B} - 0\right]$$
$$= 1 - \frac{t_A}{2t_B}$$

So the probability of A arriving first is $1 - \frac{t_A}{2t_B}$.

ex 4) We have:

$$f_{X,Y}(x,y) = \begin{cases} c & 0 \le x \le y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

a) Find c.

So we know the total area needs to be 1, so

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c\,dx\,dy = 1$$

We want more restrictions the further in the integral is, so we can say $y \in [0,1]$ and $x \in [0,y]$:

$$\int_{0}^{1}\int_{0}^{y} c\,dx\,dy = 1$$

$$\int_{0}^{1} cx\Big|_0^y\,dy = 1$$

$$\int_{0}^{1} cy\,dy = 1$$

$$\frac{cy^2}{2}\Big|_0^1 = 1$$

$$\frac{c}{2} = 1$$

$$c = 2$$

b) Find the marginal pdfs: $f_X(x)$ and $f_Y(y)$

Remember that we can find these by summing over all possible values of the other variable.

$$\therefore\ f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

Where our lower bound on y is x, and the upper bound is 1.

$$= \int_{x}^{1} 2\,dy$$

$$= 2y \Big|_x^1$$
$$= 2 - 2x,$$

on the bounds of x: $[0,1]$.

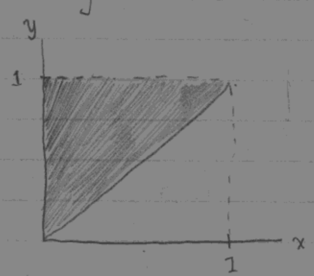
$$\therefore \quad f_X(x) = \begin{cases} 2-2x, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Let's do the same for $f_Y(y)$.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$
$$= \int_0^y 2\,dx$$
$$= 2x \Big|_0^y$$
$$= \begin{cases} 2y, & 0 \le y \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

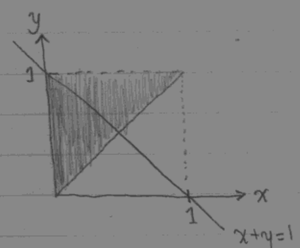
c) The probability that $X + Y < 1$.

Let's look back at our original pdf: its bounds are $0 \le x \le y \le 1$.



The shaded area represents the valid values of x and y for our function.

Remember that integrating over this entire area gives us an overall probability of 1. $X + Y < 1$ means we want all values of x and y such that $x + y < 1$. So what area does that mean?



To the left is the same plot with one difference – I've drawn in the line where $x+y=1$.

We want $x+y<1$, so we actually want all valid values BELOW this line, which is exactly half of the valid integration region

Since our pdf : $\int_y \int_x 2 \, dx \, dy$ is a uniform distribution, taking out half of the valid values yields us exactly half of the final value.

$$\int_0^1 \int_0^y 2 \, dx \, dy = 1$$

$\downarrow$ exactly half

$= 1/2$

And as such the probability of $X+Y<1$ is $1/2$.

ex 10) Suppose we have $X_1 \dots X_5$ that are independent, and have the same continuous distribution. Show that

$$P(X_3 < X_5 < X_1 < X_4 < X_2) = \frac{1}{5!}$$

Since the distribution is continuous, there are an infinite number of possible results, so we can effectively argue that each X_i has a distinct value.

Now this reduces to an "n boxes, n values" problem, like one we'd see in chapters 1/2

$$\square < \square < \square < \square < \square$$

$X_1 \quad X_2 \quad X_3 \quad X_4 \quad X_5$

There are 5 ways to fill box 1, 4 to fill box 2, etc, giving us 5! permutations. Each of these permutations is equally likely to occur, meaning the specific permutation

$$X_3 < X_5 < X_1 < X_4 < X_2$$

has a $1/5!$ chance to occur.

ex 13) Suppose X and Y have the joint density $f(x,y)$. In each case, $f(x,y) = 0$ otherwise. Are X and Y independent?

a) $f(x,y) = xe^{-x(1+y)}$ for $x,y \geq 0$

Let's solve for $f_X(x)$ and $f_Y(y)$.

$$\begin{aligned}
f_X(x) &= \int_0^\infty xe^{-x(1+y)}\,dy \\
&= \int_0^\infty xe^{-x-xy}\,dy \\
&= x\int_0^\infty \frac{e^u\,du}{-x} \quad \longleftarrow \text{ where } u = -x-xy, \quad \frac{du}{-x} = dy \\
&= -\int_0^\infty e^u\,du \\
&= -e^u \Big|_0^\infty \\
&= -e^{-x-xy}\Big|_0^\infty \\
&= -e^{-x(1-\infty)} - \left[-e^{-x(1-0)}\right] \\
&= e^{-x}
\end{aligned}$$

Now we need $f_Y(y)$.

$$f_Y(y) = \int_0^\infty x e^{-x(1+y)} \, dx$$

This one's actually pretty rough, need to integrate by parts. Kilian Miller, are you proud of me?

$$dg = fg - \int_0^\infty g \, df$$

where: $f = x \rightarrow df = dx$

$$dg = e^{-x(1+y)} dx \rightarrow g = \frac{e^{-x(1+y)}}{-1-y}$$

$$\Rightarrow = \frac{xe^{-x(1+y)}}{-1-y} - \int_0^\infty \frac{e^{-x(1+y)}}{-1-y} dx$$

$$= \frac{1}{-1-y}\left(xe^{-x(1+y)} - 1\right) - \int_0^\infty e^{-x(1+y)} dx$$

Then we integrate by substitution like before.

$$= \frac{1}{-1-y}\left(xe^{-x(1+y)} - \frac{e^{-x(1+y)}}{(-1-y)^2}\right)$$

$$= -\frac{e^{-x(y+1)}(xy+x+1)}{(y+1)^2}\Bigg|_0^\infty$$

$$= -\frac{e^{-\infty}(\infty+\infty+1)}{(y+1)^2} - \left[\frac{e^{-0}(0y+0+1)}{(y+1)^2}\right]$$

$$= \frac{1}{(y+1)^2}$$

Since $\frac{e^{-x}}{(y+1)^2} \neq xe^{-x(1+y)}$ these are not independent.

b) $f(x,y) = 6xy^2$, $x,y \geq 0$ and $x+y \leq 1$

$f_X(x)$: $y \geq 0$

$y \leq 1-x$

$$f_X(x) = \int_0^{1-x} 6xy^2 \, dy$$

$$= 6x \int_0^{1-x} y^2 \, dy$$

$$= 6x \left[\frac{y^3}{3} \right]_0^{1-x}$$

$$= 6x \left[\frac{(1-x)^3}{3} \right]$$

$$= 2x(1-x)^3$$

$f_Y(y)$: $x \geq 0$

$x \leq 1-y$

$$f_Y(y) = \int_0^{1-y} 6xy^2 \, dx$$

$$= 6y^2 \left[\frac{x^2}{2} \right]_0^{1-y}$$

$$= 3y^2(1-y)^2$$

Obviously $2x(1-x)^3 3y^2(1-y)^2 \neq 6xy^2$, so they are not independent.

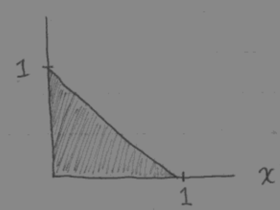
Here's one of those times where we randomly introduce theorems in the middle of the problem sets.

If $f(x,y)$ can be written as $g(x)h(y)$ and X and Y are independent, there exists a constant c such that

$$f_X(x) = c\,g(x) \qquad f_Y(y) = \frac{h(x)}{c}$$

Notice that this means the range of (x,y) MUST be a rectangle $\rightarrow$ one side being the range of x and one being the range of y.

The bounds of x and y here are $x, y \geq 0$
$x + y \leq 1$



This is not rectangular, therefore X and Y are not independent.

c) $f(x,y) = 2xy + x$

This can be rewritten as:

$$f(x,y) = 2x(y+1)$$

$$g(x) = 2x \qquad h(y) = y+1$$

where $c=1$, therefore these are independent.

ex 17) Suppose X_1 to X_m are independent and have geometric distributions with parameter p. Find $P(X_1 = k \mid X_1 + \ldots + X_m = n)$.

To jog our memory, geometric distribution is given as

$$P\{X=n\} = (1-p)^{n-1} p$$

which tells us the probability the first success occurs on the n^{th} trial.

In the "given" portion of the conditional, we have that the sum of a bunch of geometric variables is equal to some constant n.

So if $P\{X=n\}$ is the probability success 1 occurs at trial n, what does $P\{X_1+\ldots+X_m=n\}$ mean? Let's take it step by step:

$$P\{X=n\} = \underbrace{(1-p)^{n-1}}_{\substack{n-1 \\ \text{failures distributed} \\ \text{in } n-1 \text{ spots}}} \underbrace{p}_{\text{singular success}}$$

$$\therefore P\{X=3\} = (1-p)^2 p$$

Now let's add a random variable to the sum.

$$P\{X_1+X_2=3\}$$

Each one contributes a singular success to our grand scheme, so now we have two successes and 1 failure, to be distributed in 2 available spaces (as we want the 2nd success on the 3rd trial).

$$\therefore P\{X_1+X_2=3\} = \binom{2}{1}(1-p)^1 p^2$$

Now we can generalize this:

$$P\{X_1+\ldots+X_m=n\} = \binom{n-1}{m-1} p^m (1-p)^{n-m}$$

which is just the negative binomial distribution.

Then, we have $P\{X_1 = k\}$ which is just $(1-p)^{k-1}p$

$$\therefore \quad P\{X_1 = k \mid X_1 + \ldots + X_m = n\}$$

$$= \frac{P\{X_1 = k\}\, P\{X_2 + \ldots + X_m = n-k\}}{P\{X_1 + \ldots + X_m = n\}}$$

notice this is $n-k$ because X_1 is k already

So the question has become "what is the probability that the first success occurs on the k^{th} trial, given that the m^{th} success occurs on the n^{th} trial?".

$$= \frac{(1-p)^{k-1}\, p \binom{n-k-1}{m-2} (1-p)^{(n-k)-(m-1)}\, p^{m-1}}{\binom{n-1}{m-1} p^m (1-p)^{n-m}}$$

$$= \frac{\cancel{p^m}\, \cancel{(1-p)^{n-m}} \binom{n-k-1}{m-2}}{\cancel{p^m}\, \cancel{(1-p)^{n-m}} \binom{n-1}{m-1}}$$

$$= \frac{\binom{n-k-1}{m-2}}{\binom{n-1}{m-1}}$$

ex 21) Suppose X_1 and X_2 have joint density

$$f(x_1, x_2) = \begin{cases} 1 & 0 < x_1, x_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the joint density of $Y_1 = \dfrac{X_1}{X_2}$, $Y_2 = X_1 X_2$

1) Check that it can be solved:

$$X_1 = X_2 Y_1 \longrightarrow Y_2 = (X_2 Y_1) X_2 \longrightarrow X_1 = Y_1 \sqrt{\frac{Y_2}{Y_1}}$$

$$Y_2 = X_2^2 Y_1 \qquad\qquad = \sqrt{Y_1 Y_2}$$

$$X_2 = \sqrt{\frac{Y_2}{Y_1}}$$

Yes, it can

2) Check the Jacobian:

$$J\left(\sqrt{y_1 y_2}, \sqrt{\frac{y_2}{y_1}}\right):$$

$$\frac{\partial g_1}{\partial y_1} = \frac{d}{dy_1} y_1^{1/2} y_2^{1/2}$$
$$= y_1^{-1/2}(1/2) y_2^{1/2}$$

$$\frac{\partial g_2}{\partial y_2} = \frac{d}{dy_2} y_2^{1/2} y_1^{-1/2}$$
$$= y_1^{-1/2}(1/2) y_2^{-1/2}$$

$$\frac{\partial g_1}{\partial y_2} = \frac{d}{dy_2} y_1^{1/2} y_2^{1/2}$$
$$= y_1^{1/2}(1/2) y_2^{-1/2}$$

$$\frac{\partial g_2}{\partial y_1} = \frac{d}{dy_1} y_2^{1/2} y_1^{-1/2}$$
$$= y_2^{1/2}(-1/2) y_1^{-3/2}$$

$$\therefore J\left(\sqrt{y_1 y_2}, \sqrt{\frac{y_2}{y_1}}\right) = [y_1^{-1/2}(1/2) y_2^{1/2}][y_1^{-1/2}(1/2) y_2^{-1/2}] - [y_1^{1/2}(1/2) y_2^{-1/2}][y_2^{1/2}(-1/2) y_1^{-3/2}]$$
$$= \frac{1}{4y_1} + \frac{1}{4y_1}$$
$$= \frac{1}{2y_1}$$
$$\neq 0 \text{ if } y_1 \neq 0.$$

Since we've solved this Jacobian with respect to y, there's no need to take the inverse.

$$\therefore f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(x_1,x_2)\,|J(x_1,x_2)|^{-1}$$
$$= f_{X_1,X_2}(x_1,x_2)\,|J(y_1,y_2)|$$
$$= (1)\left|\frac{1}{2y_1}\right|$$
$$= \frac{1}{2y_1} \text{ where } 0 < y_1 y_2 < 1 \text{ and } 0 < y_2 < y_1$$

How did we get these bounds?

$Y_1 = X_1 X_2$ $\qquad$ $Y_2 = \frac{X_1}{X_2}$

$$\therefore \quad Y_1 Y_2 = X_1 X_2 \left(\frac{X_1}{X_2}\right)$$
$$= X_1^2$$

Since X_1^2 has bounds of $0 < X_1 < 1$,
$0 < X_1^2 < 1$,
meaning $0 < Y_1 Y_2 < 1$.

But how about $0 < Y_2 < Y_1$? Why does Y_1 need to be larger than Y_2? This one I have no clue. If someone wants to tell me where it came from, that'd be great.