Probability and Statistics for JMC
Solutions 5 — Joint Random Variables

1. Suppose the joint pdf of a pair of continuous RVs is given by

$$f(x,y) = \begin{cases} k(x+y), & 0 < x < 2,\ 0 < y < 2 \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Find the constant k.

   The pdf must be normalized: $1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy$.

   $1 = k\int_0^2\int_0^2 (x+y)dxdy = 8k$, which means $k = \frac{1}{8}$.

   (b) Find the marginal pdfs of X and Y.

   Marginal pdf for X is $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \frac{1}{8}\int_0^2 (x+y)dy = \frac{1}{8}(2x+2)$.

   So $f_X(x) = \frac{1}{4}(x+1)$ for $0 < x < 2$, and $f_X(x) = 0$ otherwise. Switching $x \leftrightarrow y$ we find $f_Y(y) = \frac{1}{4}(y+1)$ when $0 < y < 2$ and 0 otherwise.

   (c) Are X and Y independent?

   No. Independence means that $f_{XY}(x,y) = f_X(x)f_Y(y)$. This is not the case here: $\frac{1}{8}(x+y) \neq \frac{1}{4}(x+1)\frac{1}{4}(y+1)$.

2. A manufacturer has been using two different manufacturing processes to make computer memory chips. Let X and Y be two continuous random variables, where X denotes the time to failure for chips made by process A and Y denotes the time to failure for chips made by process B. Assuming that the joint pdf of (X,Y) is

$$f(x,y) = \begin{cases} (ab)e^{-(ax+by)} & x, y > 0 \\ 0 & \text{otherwise,} \end{cases}$$

   where $a = 10^{-4}$ and $b = 1.2 \times 10^{-4}$, determine $\mathrm{P}(X > Y)$.

   We need to figure out what region of the xy-plane corresponds to the event $X > Y$ and then integrate the joint pdf over this region. It is the region below the diagonal line $y = x$ (and we can ignore $x < 0$ because the pdf is 0 there). Therefore,

$$\begin{aligned}
\mathrm{P}(X > Y) &= \int_0^{\infty} dx \int_0^{x} dy\,(ab)e^{-(ax+by)} = ab\int_0^{\infty} dx e^{-ax}\int_0^{x} dy e^{-by} \\
&= ab\int_0^{\infty} dx e^{-ax}\left[-\frac{e^{-by}}{b}\right]_{y=0}^{x} = a\int_0^{\infty} dx e^{-ax}\left(1 - e^{-bx}\right) \\
&= a\int_0^{\infty} dx\left(e^{-ax} - e^{-(a+b)x}\right) = a\left[\frac{-e^{-ax}}{a} + \frac{e^{-(a+b)x}}{a+b}\right]_{x=0}^{\infty} = \frac{b}{a+b} = 0.545454.
\end{aligned}$$

3. The joint probability mass function of two discrete random variables X and Y is given by $p(x,y) = cxy$ for $x = 1,2,3$ and $y = 1,2,3$, and zero otherwise. Find

   (a) The constant c;

   The range is the set of 9 elements $(1,1),(1,2),(1,3),(2,1),\dots,(3,3)$. Normalization gives the constant: $1 = \sum_{x=1}^{3}\sum_{y=1}^{3} cxy = 36c \Rightarrow c = \frac{1}{36}$.

   (b) $\mathrm{P}(X = 2, Y = 3)$; $= c(2\cdot 3) = \frac{1}{6}$

   (c) $\mathrm{P}(X \le 2, Y \le 2)$; $= c(1\cdot 1 + 1\cdot 2 + 2\cdot 1 + 2\cdot 2) = \frac{1}{4}$

   (d) $\mathrm{P}(X \ge 2)$; $= c(2\cdot 1 + 2\cdot 2 + 2\cdot 3 + 3\cdot 1 + 3\cdot 2 + 3\cdot 3) = \frac{5}{6}$

   (e) $\mathrm{P}(Y < 2)$; $= c(1\cdot 1 + 2\cdot 1 + 3\cdot 1) = \frac{1}{6}$

   (f) $\mathrm{P}(X = 1)$; $= c(1\cdot 1 + 1\cdot 2 + 1\cdot 3) = \frac{1}{6}$

   (g) $\mathrm{P}(Y = 3)$. $= c(1\cdot 3 + 2\cdot 3 + 3\cdot 3) = \frac{1}{2}$

4. Let X and Y be continuous random variables having joint density function $f(x,y) = c\left(x^2+y^2\right)$ when $0 \le x \le 1$ and $0 \le y \le 1$, and $f(x,y) = 0$ otherwise. Determine

   (a) the constant c;

   Normalization: $1 = c\int_0^1 dx \int_0^1 dy(x^2+y^2) = c\left(\left[\frac{1}{3}x^3\Big|_0^1 + \left[\frac{1}{3}y^3\Big|_0^1\right)\right.\right. = c\frac{2}{3}$

   $\Rightarrow c = \frac{3}{2}$.

   (b) $\mathrm{P}(X < 1/2, Y > 1/2)$;

   $= c\int_0^{1/2} dx \int_{1/2}^{1} dy(x^2+y^2) = c\left(\frac{1}{2}\left[\frac{1}{3}x^3\Big|_0^{1/2} + \frac{1}{2}\left[\frac{1}{3}y^3\Big|_{1/2}^{1}\right)\right.\right. = \frac{1}{4}$

   (c) $\mathrm{P}(1/4 < X < 3/4)$;

   $= c\int_{1/4}^{3/4} dx \int_0^1 dy(x^2+y^2) = c\left(\left[\frac{1}{3}x^3\Big|_{1/4}^{3/4} + \frac{1}{2}\left[\frac{1}{3}y^3\Big|_0^1\right)\right.\right. = \frac{29}{64}$

   (d) $\mathrm{P}(Y < 1/2)$;

   $= c\int_0^1 dx \int_0^{1/2} dy(x^2+y^2) = c\left(\frac{1}{2}\left[\frac{1}{3}x^3\Big|_0^1 + \left[\frac{1}{3}y^3\Big|_0^{1/2}\right)\right.\right. = \frac{5}{16}$

   (e) whether X and Y are independent.

   Marginal pdf for X is $f_X(x) = c\int_0^1 (x^2+y^2)dy = \frac{3}{2}x^2 + \frac{1}{2}$ for $0 < x < 1$ and 0 otherwise. Similarly, $f_Y(y) = \frac{3}{2}y^2 + \frac{1}{2}$. We do not have $f(x,y) = f_X(x)f_Y(y)$ so X and Y are *not* independent.

5. If $X_1 \sim \mathrm{Gamma}(\alpha_1, \beta)$ and $X_2 \sim \mathrm{Gamma}(\alpha_2, \beta)$, with X_1 and X_2 independent, prove that $Y = X_1 + X_2 \sim \mathrm{Gamma}(\alpha_1 + \alpha_2, \beta)$.

   Use convolution theorem $f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x) f_{X_2}(y-x)dx$, where f_{X_1} and f_{X_2}

are the pdfs of Gamma RVs with corresponding parameters.

$$
\begin{aligned}
f_Y(y) &= \int_0^\infty \frac{\beta^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1-1} e^{-\beta x} \frac{\beta^{\alpha_2}}{\Gamma(\alpha_2)} (y-x)^{\alpha_2-1} e^{-\beta(y-x)} dx \\
&= \frac{\beta^{\alpha_1}\beta^{\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\beta y} \int_0^\infty x^{\alpha_1-1}(y-x)^{\alpha_2-1} dx.
\end{aligned}
$$

We can get rid of the y-dependence of the integral by a change of variables: $u = x/y$, $du = dx/y$,

$$
f_Y(y) = \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\beta y} y^{\alpha_1+\alpha_2-1} \int_0^\infty u^{\alpha_1-1}(1-u)^{\alpha_2-1} du.
$$

This is equal to the pdf of a Gamma$(\alpha_1+\alpha_2, \beta)$ random variable times a constant independent of y. Since $f_Y(y)$ is guaranteed to be normalized and Y has the same range as a Gamma$(\alpha_1+\alpha_2, \beta)$ RV, this constant must be 1.

6. Let X and Y be independent exponential random variables with parameter 1. Find the joint density function of $U = X+Y$ and $V = X/(X+Y)$, and deduce that V is uniformly distributed on $[0,1]$.

First figure out the range of the joint RVs (U, V). X and Y are each positive real numbers. Therefore, U can be any positive number and V can be anything between 0 and 1.

Second, the transformation is one-to-one since we can invert the transformation to find X and Y in terms of U and V: $X = UV$, $Y = U(1-V)$.

Now we can write down the joint density for U and V:

$$
f_{UV}(u,v) = f_{XY}(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|,
$$

where $|J| = \left|\frac{\partial(x,y)}{\partial(u,v)}\right|$ is the absolute value of the determinant of the Jacobian of the mapping between (u,v) and (x,y) (i.e. $x = uv$ and $y = u(1-v)$),

$$
|J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ 1-v & -u \end{vmatrix} = |-uv - u(1-v)| = |-u| = u.
$$

$f_{XY}(x,y)$ is just the product of the two exponential pdfs since X and Y are independent. We just need to write x and y in terms of u and v.

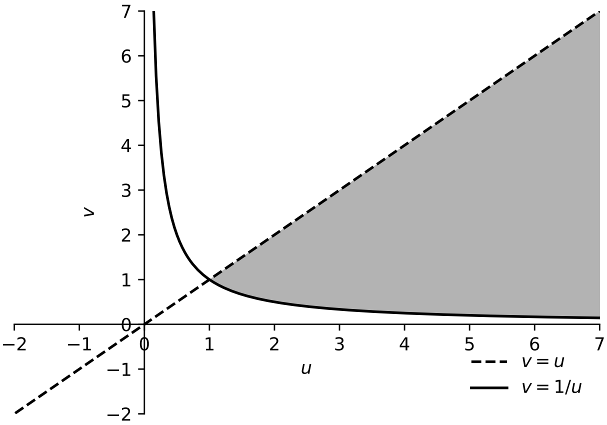
$$
\begin{aligned}
f_{UV}(u,v) &= f_X(x = uv) f_Y\left(y = u(1-v)\right) u \\
&= e^{-uv} e^{-u(1-v)} u = ue^{-u}.
\end{aligned}
$$

Since the joint density has no v dependence we know that the marginal distribution of V is uniform within its range $(0,1)$. To be explicit, if we integrated the joint density over u to get the marginal density of V we would be left with the constant 1. (Even if we don't actually compute this u-integral we know the constant must be 1 since the marginal density of V has to be a normalized pdf when integrated over the range of V.)

7. X and Y have the joint density function $f(x,y) = 1/(x^2y^2)$ when $x \geq 1$ and $y \geq 1$, and $f(x,y) = 0$ elsewhere.

(a) Compute the joint density function of U, V, where $U = XY$ and $V = X/Y$.

Step 1 is to find the range of U and V. We translate the range $x \geq 1$ and $y \geq 1$ into the range for U and V. Clearly U and V must be in the upper right quadrant of the (u,v)-plane as they are the product and quotient of positive numbers and U can take any value greater than 1. Invert the transformation to write $x = \sqrt{uv}$ and $y = \sqrt{u/v}$ (where we must take the positive root in both cases, which also shows that this is a one-to-one mapping). Then the constraints $x \geq 1$, $y \geq 1$ become $\sqrt{uv} \geq 1$ and $\sqrt{u/v} \geq 1$, which can be rearranged to $v \geq 1/u$ and $v \leq u$. The figure shows the range in the (u,v)-plane. The next step is to find the Jacobian



of the mapping.

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2}\sqrt{\frac{v}{u}} & \frac{1}{2}\sqrt{\frac{u}{v}} \\ \frac{1}{2}\frac{1}{\sqrt{uv}} & -\frac{1}{2}\sqrt{\frac{u}{v^3}} \end{vmatrix} = \left| -\frac{1}{4v} - \frac{1}{4v} \right| = \frac{1}{2v}.$$

Finally, the joint pdf of U and V is

$$f_{UV}(u,v) = f_{XY}(x,y)|J| = \frac{1}{(uv)(u/v)}\frac{1}{2v} = \frac{1}{2u^2v},$$

when $u \geq 1$ and $1/u \leq v \leq u$, and 0 otherwise.

(b) What are the marginal densities of U and V?

$$f_U(u) = \int_{-\infty}^{\infty} f_{UV}(u,v)dv = \int_{1/u}^{u} \frac{1}{2u^2v}dv = \frac{1}{2u^2}\left(\log u - \log\frac{1}{u}\right) = \frac{\log u}{u^2},$$

for $u \geq 1$ and $f_U(u) = 0$ for $u < 1$.

For V, the lower limit of the integral over u depends on whether v is greater or less than 1 (see figure).

$$f_V(v) = \int_{-\infty}^{\infty} f_{UV}(u,v)du = \begin{cases} 0 & v \leq 0 \\ \int_{1/v}^{\infty} \frac{1}{2u^2v}du = -\frac{1}{2v}u^{-1}\Big|_{1/v}^{\infty} = \frac{1}{2} & 0 < v \leq 1 \\ \int_{v}^{\infty} \frac{1}{2u^2v}du = \frac{1}{2v^2} & v > 1 \end{cases}$$

8. Prove the Law of Total Expectation: if X and Y are two random variables then

$$\mathrm{E}(X) = \mathrm{E}\left(\mathrm{E}(X \mid Y)\right),$$

where $\mathrm{E}(X \mid Y)$ is the conditional expectation of X given Y (and should be thought of as a function of the random variable Y), and the outer expectation is with respect to the marginal distribution of Y.

We'll show it for continuous RVs (the discrete case goes exactly the same way).

$$\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f_{XY}(x,y) dy \right) dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X|Y}(x|y) f_Y(y) dy dx \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx \right) f_Y(y) dy = \int_{-\infty}^{\infty} \mathrm{E}(X \mid Y = y) f_Y(y) dy = \mathrm{E}(\mathrm{E}(X \mid Y)).
\end{aligned}$$

Note, in the last integral $\mathrm{E}(X \mid Y = y)$ is just a normal function of a real number y, call it $g(y)$ and that last integral is exactly what it means to take the expected value of the random variable $g(Y)$.

9. Prove the Law of Total Variance: if X and Y are two random variables then

$$\mathrm{Var}(X) = \mathrm{E}\left(\mathrm{Var}(X \mid Y)\right) + \mathrm{Var}\left(\mathrm{E}(X \mid Y)\right),$$

where $\mathrm{Var}(X \mid Y)$ is the conditional variance of X given Y (i.e. it is the variance of X conditioned on $Y = y$, and is considered to be a function of the random variable Y).

The strategy is to begin with the definition $\mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}(X))^2]$ and then start conditioning everything on Y using the law of total expectation.

It will be clearer if we are explicit about which quantities are RVs and which are just regular numbers. Let $\mathrm{E}(X) = \mu$, which is some real number, not an RV. Let $\mathrm{E}(X|Y) = \mu_{|Y}$ to remind us that the conditional expectation of X given Y is a function of the RV Y, and thus an RV itself (but it is not a function of X). The law of total expectation says that $\mathrm{E}(\mu_{|Y}) = \mu$.

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}\Big[(X-\mu)^2\Big] = \mathrm{E}\Big[(X - \mu_{|Y} + \mu_{|Y} - \mu)^2\Big] \\
&= \mathrm{E}\Big[(X-\mu_{|Y})^2 + (\mu_{|Y}-\mu)^2 + 2(X-\mu_{|Y})(\mu_{|Y}-\mu)\Big] \\
&= \mathrm{E}\Big[(X-\mu_{|Y})^2\Big] + \mathrm{E}\Big[(\mu_{|Y}-\mu)^2\Big] + \mathrm{E}\Big[2(X-\mu_{|Y})(\mu_{|Y}-\mu)\Big].
\end{aligned}$$

The middle term is just the variance of the RV $\mu_{|Y}$, i.e. it is $\mathrm{Var}[\mathrm{E}(X|Y)]$. Use the law of total expectation on the first term to write it as $\mathrm{E}\Big[\mathrm{E}((X-\mu_{|Y})^2|Y)\Big]$. When we condition on a particular value $Y = y$ in the inner expectation, $\mu_{|Y}$ becomes just a normal number. And the conditional expectation of X given Y is that number $\mu_{|Y}$. So the inner expectation is exactly the conditional variance of X given Y and the whole first term is $\mathrm{E}\Big[\mathrm{Var}(X|Y)\Big]$. The last term is zero. To see that write it with the law of total expectation and notice that the second factor $(\mu_{|Y} - \mu)$ is just a regular number when conditioned on Y so it can be pulled outside the inner expectation. But then the inner expectation is $\mathrm{E}(X - \mu_{|Y}|Y)$ which is zero.

10. Consider the 2-class mixture model:

$$Z \sim \text{Bernoulli}(p),$$
$$X|Z \sim \begin{cases} \text{N}(\mu_0, \sigma_0^2) & \text{if } Z = 0, \\ \text{N}(\mu_1, \sigma_1^2) & \text{if } Z = 1, \end{cases}$$

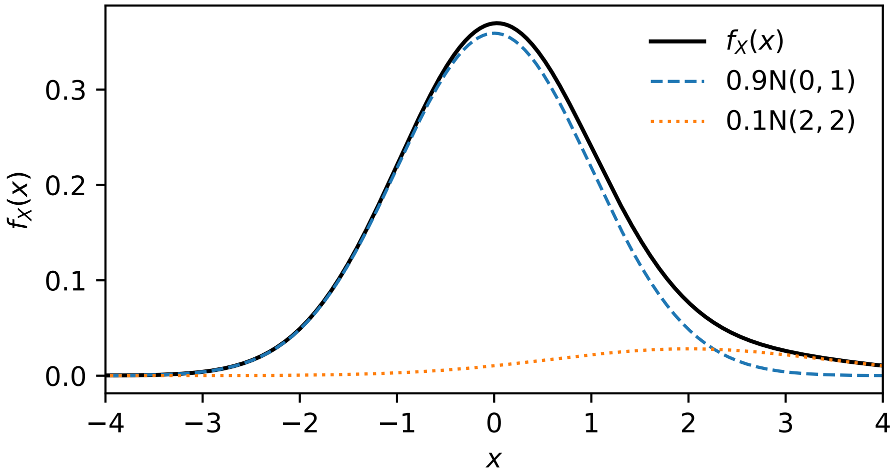
where the second line with $X|Z$ is specifying the conditional distribution of X given Z, i.e. $f_{X|Z}(x\,|\,z)$. A concrete example might be that there are two populations whose X values are distributed according to the two normal distributions. We flip a biased coin to determine Z and then, depending on whether we got heads or tails, we measure X from one or the other of the populations.

(a) Sketch the marginal pdf of X assuming the parameters $p = 0.1, \mu_0 = 0, \sigma_0^2 = 1, \mu_1 = 2, \sigma_1^2 = 2$.

Here we have a mixed discrete and continuous joint RV (Z, X). To get the marginal distribution of X we just do the obvious thing and sum the joint distribution over the possible values of Z (where the joint distribution is the product of the marginal distribution for Z and the conditional distribution of X given Z):

$$f_X(x) = \sum_{z\in\{0,1\}} f_{X|Z}(x|z)p_Z(z) = \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}}(1-p) + \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} p.$$

The marginal pdf looks like the figure below. Note the slight asymmetry due to the mixture of the two components.



(b) What are the mean and variance of X? [Hint: you can either find the marginal distribution of X first or use the laws of total expectation and total variance.]

Law of total expectation:
$\text{E}(X|Z) = \mu_0$ when $Z = 0$ and $= \mu_1$ when $Z = 1$. So the expectation of this with respect to Z is $\text{E}(X) = \text{E}(X|Z = 0)\text{P}(Z = 0) + \text{E}(X|Z = 1)\text{P}(Z = 1) = \mu_0(1 - p) + \mu_1 p$.

Law of total variance:
$\text{Var}(X|Z) = \sigma_0^2$ when $Z = 0$ and $= \sigma_1^2$ when $Z = 1$. The expectation of the conditional variance is therefore $\sigma_0^2(1 - p) + \sigma_1^2 p$.
The variance of the conditional expectation is $\text{E}[(\text{E}(X|Z) - \text{E}(X))^2]$, so

$$\begin{aligned}
\text{Var}(\text{E}(X|Z)) &= (\mu_0 - \text{E}(X))^2(1 - p) + (\mu_1 - \text{E}(X))^2 p \\
&= (\mu_0 - (\mu_0(1 - p) + \mu_1 p))^2(1 - p) + (\mu_1 - (\mu_0(1 - p) + \mu_1 p))^2 p \\
&= (\mu_0 - \mu_1)^2 p^2(1 - p) + (\mu_1 - \mu_0)^2(1 - p)^2 p \\
&= (\mu_0 - \mu_1)^2 p(1 - p).
\end{aligned}$$

Putting both pieces together, $\text{Var}(X) = \sigma_0^2(1-p) + \sigma_1^2 p + (\mu_0 - \mu_1)^2 p(1-p)$.

Alternatively, since we already have the marginal pdf (part a) we can just multiply it by x or x^2 and integrate to get $\text{E}(X)$ and $\text{E}(X^2)$. We don't have to actually do any integration since we know the mean and variance of normal pdfs (for the x^2 integrals we can make use of the fact that $\text{E}(X^2) = \text{Var}(X) + \text{E}(X)^2$).