Stochastic Calculus - Problem set 1 - Fall 2002

Exercise 1 - a
We have 600 blue balls and 300 red balls, therefore the probability to choose one blue ball is $\frac{600}{300+600} = \frac{2}{3}$. Since every time we return the chosen ball, the probability that the first n balls are blue is:

$$\left(\frac{2}{3}\right)^n$$

Exercise 1 - b
First we notice that N, the number of blue balls chosen before the first red one, takes its values in the non-negative integers $\{0, 1, 2, ...\}$. If $N = n$, we choose n blue balls before the first red one, each of them with the probability $\frac{2}{3}$, therefore:

$$P(N = n) = \left(\frac{2}{3}\right)^n \frac{1}{3}$$

Since for $|x| < 1$ the series $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ is uniformly convergent, so are all its derivatives, and we can take two successive derivatives to get:

$$\sum_{n=0}^{\infty} n x^n = \frac{x}{(1-x)^2}$$

and

$$\sum_{n=0}^{\infty} n^2 x^n = x\frac{1+x}{(1-x)^3}$$

If we plug in $x = \frac{2}{3}$ we obtain $E(N) = 2$ and $E(N^2) = 10$. Since $Var(N) = E(N^2) - E(N)^2$, the variance of N is $Var(N) = 6$.

Exercise 1 - c
Since N only takes discrete integer values, $P(N \leq 2) = P(N = 0) + P(N = 1) + P(N = 2) = \frac{1}{3}\left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2\right)$. But $P(N = 0|N \leq 2) = \frac{P(N=0)}{P(N\leq 2)}$, and it follows $P(N = 0|N \leq 2) = \frac{9}{19}$.

Exercise 1 - d
The probability that N is an even number is:

$$\sum_{k\geq 0} P(N = 2k) = \frac{1}{3}\sum_{k\geq 0}\left(\frac{2}{3}\right)^{2k} = \frac{3}{5}$$

Exercise 2 - b
Since "Good" and "Bad" are a partition of the universe Ω, we have $P(L) = P(L|G)P(G) + P(L|B)P(B)$. We already know that $P(L|G) = 0.7$ and $P(G) = 0.1$, but $P(L|B) = 1 - P(D|B) = 0.2$ and $P(B) = 1 - P(G) = 0.9$. From this we deduce $P(L) = 0.7 * 0.1 + 0.2 * 0.9 = 0.25$.

Exercise 2 - c
We can write $P(G \cap L)$ in two ways: $P(G \cap L) = P(G|L)P(L) = P(L|G)P(G)$ and thus $P(G|L) = P(L|G)\frac{P(G)}{P(L)} = 0.7\frac{0.1}{0.25} = 0.28$.

Exercise 3 - a
The mean of this random variable is given by:

$$E(X) = 2\int_0^1 x(1-x)dx = \frac{1}{3}$$

and from the second moment,

$$E(X^2) = 2\int_0^1 x^2(1-x)dx = \frac{1}{6}$$

we deduce the variance $Var(X) = E(X^2) - E(X)^2 = \frac{1}{18}$.

Exercise 3 - b
I see at least 2 ways to do this:
1) The characteristic function of the sum is the product of the characteristic functions, by independence, and what corresponds to a product in the Fourier space is a convolution, therefore:

$$f_Y(y) = \int_{\mathbb{R}} f_{X_1}(x) f_{X_2}(y-x)dx$$

2) It is possible to make a linear change of variables such as $Y = X_1 + X_2$ and $Z = X_2$ to find the joint density of the couple (Y, Z), then integrate over Z to get the marginal PDF for Y.
Whichever method you choose, we end up with having to calculate the following integral:

$$f_Y(y) = 4\int (1-x)(1-(y-x))\mathbf{1}_{0\le x\le 1}\mathbf{1}_{0\le y-x\le 1}dx$$

Simple algebra gives $\mathbf{1}_{0\le y-x\le 1} = \mathbf{1}_{y-1\le x\le y}$, and therefore we have to look at 2 separate cases.
1) If $0 \le y \le 1$ then $\mathbf{1}_{y-1\le x\le y}\mathbf{1}_{0\le x\le 1} = \mathbf{1}_{0\le x\le y}$ and the integral is:

$$f_Y(y) = 4\int_0^y (1-x)(1-(y-x))dx = 4\left(y(1-y) + \frac{y^3}{6}\right)$$

2) If $1 \le y \le 2$ then $\mathbf{1}_{y-1\le x\le y}\mathbf{1}_{0\le x\le 1} = \mathbf{1}_{y-1\le x\le 1}$ and the integral is:

$$f_Y(y) = 4\int_{y-1}^1 (1-x)(1-(y-x))dx = -\frac{2}{3}(y-2)^3$$

If we are in either situation $y \le 0$ or $y \ge 2$, the integral is zero.

Exercise 3 - c
By linearity of the expectation, we have $E(Y) = E(X_1 + X_2) = E(X_1) + E(X_2) = 2E(X) = \frac{2}{3}$.
Since X_1 and X_2 are independent, the variance of their sum is the sum of their variance, $Var(Y) = 2Var(X) = \frac{1}{9}$.

Exercise 3 - d
We use the expression we found in question b) to calculate

$$P(Y > 1) = -\frac{2}{3}\int_1^2 (y-2)^3 dy = \frac{1}{6}$$

Exercise 3 - e
If we denote by $S_{100} = X_1 + X_2 + ... + X_{100}$ we want to estimate $P(S_{100} > 34)$ using the Central Limit Theorem. We substract the expectation and normalize, to find the quantity we want to calculate is the same as $P(\frac{S_{100}-100E(X)}{\sigma\sqrt{100}} > \frac{34-100E(X)}{\sigma\sqrt{100}})$, where σ is the standard deviation of X.

The Central Limit Theorem says that the left handside converges in distribution to a Gaussian with mean 0 and variance 1, and according to the results of question a), we have $E(X) = \frac{1}{3}$ and $\sigma = \frac{1}{3\sqrt{2}}$. If $\Phi(x) = \frac{1}{2\pi}\int_{-\infty}^{x} e^{-\frac{z^2}{2}}dz$ is the CDF of a standard Gaussian, then the probability we want to estimate is $1 - \Phi(0.2828) = 0.389$.

Exercise 4 - a
We have:

$$P(X^2 + Y^2 \leq 1) = \frac{1}{8\pi}\iint_{x^2+y^2\leq 1}(4 - (x^2 + y^2))dxdy$$

The fact we integrate over a circular domain, and the function explicitely depends on $(x^2 + y^2)$ indicates it may be a good idea to switch to polar coordinates. If we write $x = r\cos\theta$ and $y = r\sin\theta$, then the Jacobian matrix of the transformation is $\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}$ whose determinant is r, and therefore formally we have $dxdy = rdrd\theta$. Thus we can write the above integral as:

$$P(X^2 + Y^2 \leq 1) = \frac{1}{8\pi}\int_0^1\int_0^{2\pi} r(4 - r^2)d\theta dr = \frac{7}{16}$$

Exercise 4 - b
For a given x, it follows from the constraint $x^2 + y^2 \leq 4$ that $|y| \leq \sqrt{4 - x^2}$. Therefore if for a given $x \in [-2; 2]$ we integrate with respect to y, we get:

$$f_X(x) = \frac{1}{8\pi}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}(4 - (x^2 + y^2))dy = \frac{1}{6\pi}(4 - x^2)^{\frac{3}{2}}$$

Exercise 4 - c
By definition we have $Cov(X, Y) = E(XY) - E(X)E(Y)$. The joint density, both as a function of x or as a function of y alone is even. Therefore its expectation over a symmetrical domain is 0, and we have $E(X) = E(Y) = 0$. The same kind of argument proves that this is still true for $E(XY)$, but to be convinced we can do the simple calculation:

$$E(XY) = \frac{1}{8\pi}\iint_{x^2+y^2\leq 4} xy(4 - (x^2 + y^2))dxdy = \frac{1}{8\pi}\int_0^2 r^3(4 - r^2)\int_0^{2\pi}\cos\theta\sin\theta d\theta dr$$

and the integral with respect to θ is zero.

Exercise 4 - d
If we look at the event $P(X \geq \sqrt{2})$, it is easy to see that it is zero if $Y \geq \sqrt{2}$, and it is not zero if $Y < \sqrt{2}$. Therefore X and Y are not independent. We know that the covariance of two independent random variables is zero, but except in the situation where the variables are jointly Gaussian, a zero covariance does not imply independence. This exercise provides us with a concrete example.

Exercise 4 - e
If $U = X^2$ and $V = Y^2$, then $X = \sqrt{U}$ and $Y = \sqrt{V}$, and the determinant of the Jacobian matrix $\begin{pmatrix}\frac{1}{2\sqrt{u}} & 0\\ 0 & \frac{1}{2\sqrt{v}}\end{pmatrix}$ is $\frac{1}{4\sqrt{uv}}$. Now we have to find the image of the domain $D = \{(x, y)|x^2 + y^2 \leq 4\}$ under the map $u = x^2$,$v = y^2$. It is obvious that $u \geq 0$, $v \geq 0$, and $u + v \leq 4$, and that every point of this triangle T has a preimage in the original disk. As a matter of the fact, except the origin, each point has 4 preimages: we have a onto map that is not into. In order to make our change of variables, we need to find a C^1-diffeomorphism. But if we restrict our original domain to any quarter of the disk, we have a one-to-one map and we can apply our theorem. Doing so on each quarter gives the same result, therefore in the resulting density, a factor 4 comes out of the disk

decomposition. If we take any bounded continuous function f, we then have

$$E(f(U,V)) = \frac{1}{8\pi}\iint_D f(x^2,y^2)(4-(x^2+y^2))dxdy = 4\frac{1}{8\pi}\iint_T f(u,v)\frac{4-(u+v)}{4\sqrt{uv}}dudv$$

But if this is true for any bounded continuous function f, then we must have:

$$f_{(U,V)}(u,v) = \frac{1}{8\pi}\frac{4-(u+v)}{\sqrt{uv}}\mathbf{1}_{(u,v)\in T}$$

Exercise 4 - f

This is just a calculation, and I am not going to write every step of it. First we need to calculate

$$E(X^2) = \frac{1}{8\pi}\int_0^4\int_0^{4-u} u\frac{4-(u+v)}{\sqrt{uv}} = \frac{2}{3}$$

and by symmetry we have $E(Y^2) = \frac{2}{3}$ as well. Now we can switch to polar coordinates to calculate the integral:

$$E(X^2Y^2) = \frac{1}{8\pi}\iint_{x^2+y^2\leq 4} x^2y^2(4-(x^2+y^2))dxdy = \frac{1}{3}$$

It follows that $Cov(X^2,Y^2) = E(X^2Y^2) - E(X^2)E(Y^2) = -\frac{1}{9}$. If X and Y were independent, X^2 and Y^2 would be independent as well, and we would have $Cov(X^2,Y^2) = 0$. The fact it is not zero gives another proof of their non independence.