Recitation Session 2 Solutions

1. *Triangle Island.* See lecture notes for the full solution. *Hint*: If you sum over horizontal infinitesimal rectangular strips the order of integration is $dx\,dy$, and vica versa.

2. *Sums.*

a. When $X = x$ we can identify $Y = z - x$, but this is not the only case where we have the sum equal to z, we have to sum over all disjoint cases that gives the desired sum.

$$\mathrm{P}\left(X + Y = z\right) = \sum_{\text{all } x} \mathrm{P}\left(X = x, Y = z - x\right) \tag{1}$$

b. Draw the (x, y)-plane and look at the thin strip that describes $Z \in dz$, it is an infinitely long infinitesimally thick strip with slope -1. Then the density of Z will be $f_Z(z) = P(Z \in dz)/dz$. Let's break that into small chunks in the x axis, over $(x, x + dx)$.

$$\mathrm{P}\left(X \in dx, Z \in dz\right) = f(x, z - x)dxdz \tag{2}$$

This gives the joint density of X and Z. To get the marginal density of Z integrate over x.

$$\mathrm{P}\left(Z \in dz\right) = \Big[\int_{-\infty}^{\infty} f(x, z - x)dx\Big]dz \tag{3}$$

This is as if you are summing over infinitesimal parallelograms. If X and Y are independent then their density factors out to give us the *convolution* formula:

$$f_Z(z) = \int_{-\infty}^{\infty} f(x)f(z - x)dx \tag{4}$$

c. Note that there are two separate cases to consider: $0 < z < 1$ and $1 < z < 2$. Look at the event $\{X + Y = z\}$. The area is $z\,dz + 1/2(dz)^2$, here the second term is negligible, so simply: $\mathrm{P}\left(Z \in dz\right) = z\,dz$. For the second region it is $\mathrm{P}\left(Z \in dz\right) = (2 - z)\,dz$. As an excercise draw the densities of sums of three, and four uniform $(0, 1)$ variables. One way to draw them is to draw the (X, Y) scatter density and take projections onto any one direction that you would like to compute: e.g. horizontal projection to compute the marginal density of Y, or diagonal (along the constants of $X + Y$) to compute the sum. Then draw $X + Y$ on one axis and W (the third uniform) on another axis and repeat the same. How does the shape of the distribution change as you keep adding uniform random variables?

3. *Discrete uniform - binomial.*

a. Let P be the proportion of black balls in the box picked. Possible values for P are $\{0, \frac{1}{N}, \frac{1}{N}, \dots, \frac{N-1}{N}, 1\}$ once a box is picked, so that P is known, the distribution of S_n is binomial(n, p):

$$\mathrm{P}\left(S_n = k | P = p\right) = \binom{n}{k} p^k (1 - p)^{n-k} \tag{5}$$

All $N+1$ possible values for p are equally likely because the priors are uniform, therefore one can average over those values (or just write the definition of conditional probability), and substitute $p = b/N$ to get the unconditional density of S_n:

$$\mathrm{P}\left(S_n = k\right) = \sum_p \binom{n}{k} p^k (1-p)^{n-k} \frac{1}{N+1} \tag{6}$$

$$= \binom{n}{k} \frac{1}{(N+1)N^n} \sum_{b=0}^{N} b^k (N-b)^{n-k} \tag{7}$$

Evaluate this for all k when $N = 2$ and $N = 3$.

b. Recognize this as the discrete approximation to the beta integral. $B(k+1, n-k+1) = \int_0^1 p^k (1-p)^{n-k} dp$ then use the expression of the beta integral in terms of gamma functions, as well as $\Gamma m + 1 = m!$ if m is integer: $B(k+1, n-k+1) = \binom{n}{k}^{-1} \frac{1}{n+1}$. Hence as $N \to \infty$ the distribution is the uniform distribution on $0, 1, \ldots, n$:

$$\mathrm{P}\left(S_n\right) \to \binom{n}{k} \binom{n}{k}^{-1} \frac{1}{n+1} = \frac{1}{n+1} \tag{8}$$

Remark: Here don't worry about the formulas for the beta integral, and the gamma function. It is much more important to understand why the result makes sense. What would you guess without any calculation? Why does uniform distribution makes sense?

4. *Continuous uniform - binomial.*

a. First try to guess this by looking at the answer to the previous question. What does the prior resemble when N tends to infinity? It looks more and more like the uniform distribution over $(0, 1)$ which is what we are asked now. So the answer should be $\frac{1}{n+1}$, let's calculate it:

$$\mathrm{P}\left(S_n = k\right) = \int \mathrm{P}\left(S_n = k | U = p\right) f(p)\, dp \tag{9}$$

$$= \int_0^1 \binom{n}{k} p^k (1-p)^k dp = \frac{1}{n+1} \tag{10}$$

$$\tag{11}$$

b. By the Bayes rule:

$$\mathrm{P}\left(U \in dp | S_n = k\right) = \frac{\mathrm{P}\left(S_n = k | U \in dp\right) \mathrm{P}\left(U \in dp\right)}{\mathrm{P}\left(S_n = k\right)} \tag{12}$$

$$= (n+1) \binom{n}{k} p^k (1-p)^{n-l} dp \tag{13}$$

Conlcusion: the distribution of U given $S_n = k$ is beta$(k+1, n-k+1)$.

c. If we are given both $U = p$ and $S_n = k$ then $\mathrm{P}\left(\text{next trial is success}\right)$ is just p because trials are independent. If we are given only $S_n = k$ we don't know what U is (it is a r.v. whose distribution is found in the previous part). Then using the mean for beta distribution we get $\frac{k+1}{n+2}$. Again, here don't worry about the details of beta integral. Also the concept of mean will be covered in the upcoming lecture. But it is important to think about its meaning. If $k = n$, the probability of $n + 1^{st}$ success given n successes is $(n + 1)/(n + 2)$.

Discussion by Jim Pitman: "This formula, for the probability of one more success given a run of n successes in independent trials with unknown success probability assumed uniformly distributed on $(0, 1)$, is known as Laplace's law of succession. Laplace illustrated his formula by calculating the probability that the sun will rise tomorrow, given that it has risen daily for 5000 years, or $n = 1,826,213$ days. But this kind of application is of doubtful value. Both the assumption of independent trials with unknown p and the uniform prior distribution of p make little sense in this context."

5. *Sampling from the unit disc.*

   a. In this type of problems one can either first find the pdf and derive cdf or first find the cdf and derive the pdf. For the first part it is easier to find the density first. Look at the event $(X \in dx)$ for small dx (again recall the slight abuse of notation this in fact means for a point x we look at $\{w|X(w) \in (x, x + dx)\}$). Then the area of that rectangle is $2\sqrt{a - x^2}dx$, divide this by the total area π to get the pdf. And integrate the pdf to get the cdf, you can leave in the integral form, it is not easy to evaluate.

   b. Consider the event $(R \in dr)$ it is an annulus inside the disc.

$$\mathrm{P}\left(R \in dr\right) = \frac{\text{Area of annulus from } r \text{ to } r + dr}{\text{Total area}} = \frac{\pi(r + dr)^2 - \pi(r)^2}{\pi} = 2r\, dr \tag{14}$$

6. *Maximum of independent random variables.*

   a. When dealing with extreme values, using cdf is in general the easier approach.

$$\begin{aligned}
F_{\max}(x) &= \mathrm{P}\left(X_{\max}\right) && (15)\\
&= \mathrm{P}\left(X_1 \le x, X_2 \le x, \ldots, X_n \le x\right) && (16)\\
&= \mathrm{P}\left(X_1 \le x\right)\mathrm{P}\left(X_2 \le x\right)\ldots\mathrm{P}\left(X_n \le x\right) && (17)\\
&= F_1(x)F_2(x)\ldots F_n(x) && (18)
\end{aligned}$$

   b. For this part find the cdf of $X_{\min}$:

$$F_{\min}(x) = 1 - (1 - F_1(x))(1 - F_2(x))\ldots(1 - F_n(x)) \tag{19}$$

   Then find cdf of $\exp(\lambda_i)$ which is $1 - e^{-\lambda_i x}$ for x non-negative. Apply the formula to get

$$F_{\min}(x) = 1 - e^{-(\lambda_1+\lambda_2+\cdots+\lambda_n)x} \tag{20}$$

   which is the cdf of exponential distribution with rate $\lambda_1 + \lambda_2 + \cdots + \lambda_n$. Therefore sum of exponential distributions is still exponential with the rate being equal to the sum of individual rates.