CS 70	Discrete Mathematics and Probability Theory	
Spring 2019	Babak Ayazifar and Satish Rao	HW 11

1 Family Planning

Mr. and Mrs. Brown decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Browns have. Let C be the total number of children they have.

(a) Determine the sample space, along with the probability of each sample point.

(b) Compute the joint distribution of G and C. Fill in the table below.

	$C=1$	$C=2$	$C=3$
$G=0$			
$G=1$			

(c) Use the joint distribution to compute the marginal distributions of G and C and confirm that the values are as you'd expect. Fill in the tables below.

$\mathbb{P}(G=0)$	
$\mathbb{P}(G=1)$	

$\mathbb{P}(C=1)$	$\mathbb{P}(C=2)$	$\mathbb{P}(C=3)$

(d) Are G and C independent?

(e) What is the expected number of girls the Browns will have? What is the expected number of children that the Browns will have?

Solution:

(a) The sample space is the set of all possible sequences of children that the Browns can have: $\Omega = \{g, bg, bbg, bbb\}$. The probabilities of these sample points are:

$$\mathbb{P}(g) = \frac{1}{2}$$
$$\mathbb{P}(bg) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$
$$\mathbb{P}(bbg) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$
$$\mathbb{P}(bbb) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

(b)

	$C=1$	$C=2$	$C=3$
$G=0$	0	0	$\mathbb{P}(bbb)=1/8$
$G=1$	$\mathbb{P}(g)=1/2$	$\mathbb{P}(bg)=1/4$	$\mathbb{P}(bbg)=1/8$

(c) Marginal distribution for G:

$$\mathbb{P}(G=0)=0+0+\frac{1}{8}=\frac{1}{8}$$
$$\mathbb{P}(G=1)=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}=\frac{7}{8}$$

Marginal distribution for C:

$$\mathbb{P}(C=1)=0+\frac{1}{2}=\frac{1}{2}$$
$$\mathbb{P}(C=2)=0+\frac{1}{4}=\frac{1}{4}$$
$$\mathbb{P}(C=3)=\frac{1}{8}+\frac{1}{8}=\frac{1}{4}$$

(d) No, G and C are not independent. If two random variables are independent, then

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

To show this dependence, consider an entry in the joint distribution table, such as $\mathbb{P}(G=0,C=3)=1/8$. This is not equal to $\mathbb{P}(G=0)\mathbb{P}(C=3)=(1/8)\cdot(1/4)=1/32$, so the random variables are not independent.

(e) We can apply the definition of expectation directly for this problem, since we've computed the marginal distribution for both random variables.

$$\mathbb{E}(G)=0\cdot\mathbb{P}(G=0)+1\cdot\mathbb{P}(G=1)=1\cdot\frac{7}{8}=\frac{7}{8}$$
$$\mathbb{E}(C)=1\cdot\mathbb{P}(C=1)+2\cdot\mathbb{P}(C=2)+3\cdot\mathbb{P}(C=3)=1\cdot\frac{1}{2}+2\cdot\frac{1}{4}+3\cdot\frac{1}{4}=\frac{7}{4}$$

2 More Family Planning

(a) Suppose we have a random variable $N \sim Geom(1/3)$ representing the number of children of a randomly chosen family. Assume that within the family, children are equally likely to be boys and girls. Let B be the number of boys and G the number of girls in the family. What is the joint probability distribution of B, G?

(b) Given that we know there are 0 girls in the family, what is the most likely number of boys in the family?

(c) Now let X and Y be independent random variables representing the number of children in two independently, randomly chosen families. Suppose that $X \sim Geom(p)$ and $Y \sim Geom(q)$. Find $\mathbb{P}(X < Y)$, the probability that the number of children in the first family (X) is less than the number of children in the second family (Y). (You may use the convergence formula for a Geometric Series: $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$ for $|r| < 1$)

(d) Show how you could obtain your answer from the previous part using an interpretation of the geometric distribution.

Solution:

(a) Conditional on $N = n$, we see that $B \sim Bin(n, \frac{1}{2})$. Then, because we have that $N = B + G$,

$$\begin{aligned}
\mathbb{P}[B = b, G = g] &= \mathbb{P}[B = b, N = b + g] \\
&= \mathbb{P}[N = b + g]\mathbb{P}[B = b | N = b + g] \\
&= (1/3)(2/3)^{b+g-1}\binom{b+g}{b}(1/2)^{b+g}
\end{aligned}$$

(b) We have that

$$\mathbb{P}[B = b | G = 0] = \frac{\mathbb{P}[B = b, G = 0]}{\mathbb{P}[G = 0]} = \frac{(1/3)(2/3)^{b-1}\binom{b}{b}(1/2)^b}{\mathbb{P}[G = 0]} = \frac{2(1/3)^b}{\mathbb{P}[G = 0]}$$

Since this decreases with b, this is maximized when $B = 1$.

Note that B can not be 0 in this case, because this would imply that $N = 0$. This is not possible since N follows a Geometric Distribution, which takes on values $1, 2, 3, \ldots$.

(c) We have that for a geometric random variable Y with parameter q, $\mathbb{P}[Y > n] = (1-q)^n$. Now, since X and Y are independent, we have that $\mathbb{P}[X = n, Y > n] = \mathbb{P}[X = n]\mathbb{P}[Y > n] = p(1-p)^{n-1}(1-q)^n$. Then:

$$\begin{aligned}
\mathbb{P}[X < Y] &= \sum_{n=1}^{\infty} \mathbb{P}[X = n, Y > n] \\
&= \sum_{n=1}^{\infty} p(1-p)^{n-1}(1-q)^n \\
&= p(1-q)\sum_{n=1}^{\infty}(1-p)^{n-1}(1-q)^{n-1} \\
&= p(1-q)\sum_{n=0}^{\infty}(1-p)^n(1-q)^n \\
&= p(1-q)\sum_{n=0}^{\infty}[(1-p)(1-q)]^n
\end{aligned}$$

$$= p(1-q) \cdot \frac{1}{1-(1-p)(1-q)}$$

$$= \frac{p(1-q)}{p+q-pq}$$

(d) If we treat the scenario of two geometric distributions as carrying out two different experiments at the same time (with probabilities p and q of success), we see that finding $\mathbb{P}[Y > X]$ is the same as seeing that Y took more trials for a success than X. Then, we can ignore all trials where both experiments failed (i.e. only consider the first success in either experiment). From this, conditioned on the probability there was indeed one success among the two experiments (which has probability $p+q-pq$ by inclusion-exclusion), the probability that $Y > X$ is just the probability that X succeeded and Y did not, which is $p(1-q)$. Then, $\mathbb{P}[Y > X] = \frac{p(1-q)}{p+q-pq}$

We could alternatively use a recursive definition of the probability that X is less than Y to arrive at the same solution. We let $Z = \mathbb{P}[X < Y]$. Then, on the first trial, for X to be less than Y, either X is successful and Y is not (with probability $p(1-q)$), or both are not successful and we repeat the experiment (with probability $p(1-q)Z$). Then, we get the equation $Z = p(1-q) + (1-p)(1-q)Z$, and we solve for Z to get $Z = \frac{p(1-q)}{p+q-pq}$, the desired answer.

3 Combining Distributions

(a) Let $X \sim Pois(\lambda), Y \sim Pois(\mu)$ be independent. Prove that $X+Y \sim Pois(\lambda + \mu)$.

Hint: Recall the binomial theorem, which states that

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$

(b) Let X and Y be defined as in the previous part. Prove that the distribution of X conditional on $X+Y$ is a binomial distribution, e.g. that $X|X+Y$ is binomial. What are the parameters of the binomial distribution?

Hint: Your result from the previous part will be helpful.

Solution:

(a) If we want $X+Y$ to take on the value k, we can have X take on any value from 0 to k as long as $Y = k - X$. Since the events of these happening are disjoint, we can write

$$\mathbb{P}[X+Y=k] = \sum_{x=0}^{k} \mathbb{P}[X = x \cap Y = k-x]$$

Since X and Y are independent, this becomes

$$\begin{aligned}\mathbb{P}[X+Y=k] &= \sum_{x=0}^{k} \mathbb{P}[X=x]\cdot\mathbb{P}[Y=k-x] \\ &= \sum_{x=0}^{k} \frac{\lambda^x e^{-\lambda}}{x!}\cdot\frac{\mu^{k-x}e^{-\mu}}{(k-x)!} \\ &= e^{-\lambda-\mu}\sum_{x=0}^{k}\frac{\lambda^x\mu^{k-x}}{x!(x-k)!}\end{aligned}$$

If we then factor out a $\frac{1}{k!}$ from each term of this summation, we get

$$\begin{aligned}\mathbb{P}[X+Y=k] &= \frac{e^{-(\lambda+\mu)}}{k!}\sum_{x=0}^{k}\lambda^x\mu^{k-x}\frac{k!}{x!(k-x)!} \\ &= \frac{e^{-(\lambda+\mu)}}{k!}\sum_{x=0}^{k}\lambda^x\mu^{k-x}\binom{k}{x}\end{aligned}$$

Applying the binomial theorem, we can replace the summation by $(\lambda+\mu)^k$, so we have

$$\mathbb{P}[X+Y=k] = \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^k}{k!}$$

Thus, the PMF of $X+Y$ matches that of a Poisson random variable with parameter $\lambda+\mu$. Since the PMF uniquely defines the distribution, this means that $X+Y \sim Pois(\lambda+\mu)$.

(b)

$$\begin{aligned}P(X=k|X+Y=n) &= \frac{P(X=k\cap X+Y=n)}{P(X+Y=n)} \\ &= \frac{\frac{\lambda^k e^{-\lambda}}{k!}\times\frac{\mu^{n-k}e^{-\mu}}{(n-k)!}}{\frac{(\lambda+\mu)^n e^{-(\lambda+\mu)}}{n!}} \\ &= \frac{n!}{k!(n-k)!}\times\frac{e^{-\lambda}e^{-\mu}}{e^{-(\lambda+\mu)}}\times\frac{\lambda^k\mu^{n-k}}{(\lambda+\mu)^n} \\ &= \binom{n}{k}(\frac{\lambda}{\lambda+\mu})^k(\frac{\mu}{\lambda+\mu})^{n-k} \\ &= \binom{n}{k}(\frac{\lambda}{\lambda+\mu})^k(1-\frac{\lambda}{\lambda+\mu})^{n-k}\end{aligned}$$

Recall that if $X=k$ and $X+Y=n$, then $Y=n-k$.

4 Darts

Yiming is playing darts. Her aim follows an exponential distribution with parameter 1; that is, the probability density that the dart is x distance from the center is $f_X(x)=\exp(-x)$. The board's radius is 4 units.

(a) What is the probability the dart will stay within the board?

(b) Say you know Yiming made it on the board. What is the probability she is within 1 unit from the center?

(c) If Yiming is within 1 unit from the center, she scores 4 points, if she is within 2 units, she scores 3, etc. In other words, Yiming scores $\lfloor 5-x \rfloor$, where x is the distance from the center. (This implies that Yimig scores 0 points if she throws it off the board). What is Yiming's expected score after one throw?

Solution:

(a) The CDF of an exponential is $\mathbb{P}[X \le x] = 1-\exp(-x)$. Therefore,

$$\mathbb{P}[X \le 4] = 1-\exp(-4).$$

(b) We are given that the dart must be within the board, which means that the dart is at least 4 units away from the center. We can use the definition of conditional probability:

$$\mathbb{P}[X \le 1 \mid X \le 4] = \frac{\mathbb{P}[X \le 1 \cap X \le 4]}{\mathbb{P}[X \le 4]} = \frac{\mathbb{P}[X \le 1]}{\mathbb{P}[X \le 4]} = \frac{1-\exp(-1)}{1-\exp(-4)}.$$

(c)

$$\begin{aligned}\mathbb{E}[\text{score}] &= \int_0^1 4\exp(-x)\,\mathrm{d}x + \int_1^2 3\exp(-x)\,\mathrm{d}x + \int_2^3 2\exp(-x)\,\mathrm{d}x + \int_3^4 \exp(-x)\,\mathrm{d}x \\ &= 4\big(-\exp(-1)+1\big) + 3\big(-\exp(-2)+\exp(-1)\big) + 2\big(-\exp(-3)+\exp(-2)\big) \\ &\qquad + \big(-\exp(-4)+\exp(-3)\big) \\ &= 4-\exp(-1)-\exp(-2)-\exp(-3)-\exp(-4).\end{aligned}$$

5 Uniform Means

Let $X_1, X_2, \ldots, X_n$ be n independent and identically distributed uniform random variables on the interval $[0,1]$ (where n is a positive integer).

(a) Let $Y = \min\{X_1, X_2, \ldots, X_n\}$. Find $\mathbb{E}(Y)$. [*Hint*: Use the tail sum formula, which says the expected value of a nonnegative random variable is $\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X > x)\,\mathrm{d}x$. Note that we can use the tail sum formula since $Y \ge 0$.]

(b) Let $Z = \max\{X_1, X_2, \ldots, X_n\}$. Find $\mathbb{E}(Z)$. [*Hint*: Find the CDF.]

Solution:

(a) To calculate $\mathbb{P}(Y > y)$, where $y \in [0,1]$, this means that each X_i is greater than y, for $i = 1, \ldots, n$, so $\mathbb{P}(Y > y) = (1-y)^n$. We then use the tail sum formula:

$$\mathbb{E}(Y) = \int_0^1 \mathbb{P}(Y > y)\,\mathrm{d}y = \int_0^1 (1-y)^n\,\mathrm{d}y = -\frac{1}{n+1}(1-y)^{n+1}\Big|_0^1 = \frac{1}{n+1}.$$

Alternative Solution 1:
As explained above, $\mathbb{P}[Y \le y] = 1-(1-y)^n$. This gives us the CDF, and if we take its derivative we'll get the probability density function $f(y) = n(1-y)^{n-1}$.

Then

$$\mathbb{E}(Y) = \int_0^1 y \cdot n(1-y)^{n-1}\,\mathrm{d}y.$$

Perform a u substitution, where $u = 1-y$ and $\mathrm{d}u = -\mathrm{d}y$. We see:

$$\begin{aligned}\mathbb{E}(Y) &= n \cdot \int_0^1 -(1-u)\cdot u^{n-1}\,\mathrm{d}u = n \cdot \int_0^1 (u^n - u^{n-1})\,\mathrm{d}u = n\left[\frac{u^{n+1}}{n+1} - \frac{u^n}{n}\right]_{u=0}^1 \\ &= n\left[\frac{(1-y)^{n+1}}{n+1} - \frac{(1-y)^n}{n}\right]_{y=0}^1 = n\left[0 - \left(\frac{1}{n+1} - \frac{1}{n}\right)\right] = n\left[\frac{1}{n} - \frac{1}{n+1}\right] = \frac{1}{n+1}.\end{aligned}$$

Alternative Solution 2:
Consider adding another independent uniform variable X_{n+1}. $\mathbb{P}(X_{n+1} < Y)$ is just the probability that X_{n+1} is the minimum, which is $1/(n+1)$ by symmetry since all the X_i's are identical. It so happens that because X_{n+1} is a uniform variable on [0,1], this probability is equal to $\mathbb{E}(Y)$. Let f_Y denote the PDF of Y.

$$\begin{aligned}\mathbb{P}(X_{n+1} < Y) &= \int_0^1 \mathbb{P}(X_{n+1} < y \mid Y = y) f_Y(y)\,\mathrm{d}y \\ &= \int_0^1 \mathbb{P}(X_{n+1} < y) f_Y(y)\,\mathrm{d}y && \text{(by independence)} \\ &= \int_0^1 y f_Y(y)\,\mathrm{d}y && \text{(CDF of the uniform distribution)} \\ &= \mathbb{E}(Y).\end{aligned}$$

Alternative Solution 3:
Since $X_1, \ldots, X_n$ are i.i.d., their values split the interval $[0,1]$ into $n+1$ sections, and we expect these sections to be of equal length because they are uniformly distributed. Therefore, $\mathbb{E}(Y) = 1/(n+1)$, the position of the smallest indicator.

(b) We could use the tail sum formula, but it turns out that the CDF is in a form that makes it easy to take an integral. If $Z \le z$, where $z \in [0,1]$, each X_i must be less than z, which happens with probability z, so $\mathbb{P}[Z \le z] = z^n$. This gives us the CDF, and if we take its derivative we'll get the probability density function $f(z) = nz^{n-1}$. Then

$$\mathbb{E}(Z) = \int_0^1 z \cdot nz^{n-1}\,\mathrm{d}z = \int_0^1 nz^n\,\mathrm{d}z = \left[n \cdot \frac{z^{n+1}}{n+1}\right]_{z=0}^1 = \frac{n}{n+1}.$$

Alternative Solution:
As in the previous part, add another independent uniform random variable X_{n+1}. The probability $\mathbb{P}(X_{n+1} > Z)$ is just the probability that X_{n+1} is the maximum, which is $1/(n+1)$ by symmetry.

$$\begin{aligned}\mathbb{P}(X_{n+1} > Z) &= \int_0^1 \mathbb{P}(X_{n+1} > z \mid Z = z) f_Z(z)\,\mathrm{d}z = \int_0^1 \mathbb{P}(X_{n+1} > z) f_Z(z)\,\mathrm{d}z \\ &= \int_0^1 (1-z) f_Z(z)\,\mathrm{d}z = \int_0^1 f_Z(z)\,\mathrm{d}z - \int_0^1 z f_Z(z)\,\mathrm{d}z \\ \frac{1}{n+1} &= 1 - \mathbb{E}(Z) \\ \mathbb{E}(Z) &= \frac{n}{n+1}\end{aligned}$$

Alternative Solution 2:
Since $X_1,\dots,X_n$ are i.i.d., their values split the interval $[0,1]$ into $n+1$ sections, and we expect these sections to be of equal length because they are uniformly distributed. The expectation of the smallest X_i is $1/(n+1)$, the expectation of the second smallest is $2/(n+1)$, etc. Therefore, $\mathbb{E}(Z) = n/(n+1)$, the position of the largest indicator.

Alternative Solution 3:
Let us define $Y_i = 1 - X_i$. Then, $Z = \max\{X_1, X_2, \dots, X_n\} = 1 - \min\{Y_1, Y_2, \dots, Y_n\}$. Observe that, although a function of X_i, the Y_i are also independent and identically distributed uniform random variables over $[0,1]$. Thus, we apply the previous part to find that $\mathbb{E}[\min\{Y_1, Y_2, \dots, Y_n\}] = 1/n+1$. As a result,

$$\begin{aligned}\mathbb{E}[Z] &= \mathbb{E}[1 - \min\{Y_1, Y_2, \dots, Y_n\}] \\ &= 1 = \mathbb{E}[\min\{Y_1, Y_2, \dots, Y_n\}] \\ &= 1 - \frac{1}{n+1} \\ &= \frac{n}{n+1}\end{aligned}$$

6 Moments of the Exponential Distribution

Let $X \sim \text{Exponential}(\lambda)$, where $\lambda > 0$. Show that for all positive integers k, $\mathbb{E}[X^k] = k!/\lambda^k$. [*Hint:* Integration by Parts.]

Solution:

The base case is $\mathbb{E}[X] = 1/\lambda$, which we already know. Using integration by parts,

$$\begin{aligned}\mathbb{E}[X^{k+1}] &= \int_0^\infty x^{k+1} \cdot \lambda \exp(-\lambda x)\,\mathrm{d}x = -x^{k+1} \exp(-\lambda x)\Big|_0^\infty + (k+1) \int_0^\infty x^k \exp(-\lambda x)\,\mathrm{d}x \\ &= \frac{k+1}{\lambda} \int_0^\infty x^k \cdot \lambda \exp(-\lambda x)\,\mathrm{d}x = \frac{k+1}{\lambda} \mathbb{E}[X^k] = \frac{(k+1)!}{\lambda^{k+1}}\end{aligned}$$

which proves the inductive step.

Another way to understand this result is to first note that for $t \in [0, \lambda)$,

$$\begin{aligned}
\mathbb{E}[\exp(tX)] &= \int_0^\infty \exp(tx)\lambda\exp(-\lambda x)\,\mathrm{d}x = \frac{\lambda}{\lambda - t}\int_0^\infty (\lambda - t)\exp\bigl(-(\lambda - t)x\bigr)\,\mathrm{d}x = \frac{\lambda}{\lambda - t} \\
&= \frac{1}{1 - t/\lambda} = \sum_{k=0}^\infty \frac{1}{\lambda^k} t^k
\end{aligned}$$

and also

$$\mathbb{E}[\exp(tX)] = \mathbb{E}\Bigl[\sum_{k=0}^\infty \frac{t^k X^k}{k!}\Bigr] = \sum_{k=0}^\infty \frac{\mathbb{E}[X^k]}{k!} t^k$$

by linearity of expectation, so by comparing the two expressions we find $\mathbb{E}[X^k] = k!/\lambda^k$.