CS 70 Discrete Mathematics and Probability Theory
Fall 2017 Kannan Ramchandran and Satish Rao

HW 12

Sundry

Before you start your homework, write down your team. Who else did you work with on this homework? List names and email addresses. (In case of homework party, you can also just describe the group.) How did you work on this homework? Working in groups of 3-5 will earn credit for your "Sundry" grade.

Please copy the following statement and sign next to it:

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up. (*signature here*)

1 Darts with Friends

Michelle and Alex are playing darts. Being the better player, Michelle's aim follows a uniform distribution over a circle of radius r around the center. Alex's aim follows a uniform distribution over a circle of radius $2r$ around the center.

(a) Let the distance of Michelle's throw be denoted by the random variable X and let the distance of Alex's throw be denoted by the random variable Y.

- What's the cumulative distribution function of X?
- What's the cumulative distribution function of Y?
- What's the probability density function of X?
- What's the probability density function of Y?

(b) What's the probability that Michelle's throw is closer to the center than Alex's throw? What's the probability that Alex's throw is closer to the center?

(c) What's the cumulative distribution function of $U = \min\{X,Y\}$?

(d) What's the cumulative distribution function of $V = \max\{X,Y\}$?

(e) What is the expectation of the absolute difference between Michelle's and Alex's distances from the center, that is, what is $\mathbb{E}[|X-Y|]$? [*Hint*: There are two ways of solving this part.]

Solution:

(a)
- To get the cumulative distribution function of X, we'll consider the ratio of the area where the distance to the center is less than x, compared to the entire available area. This gives us the following expression:

$$\mathbb{P}(X \le x) = \frac{\pi x^2}{\pi r^2} = \frac{x^2}{r^2}, \qquad x \in [0, r]$$

- Using the same approach as the previous part:

$$\mathbb{P}(Y \le y) = \frac{\pi y^2}{\pi \cdot 4r^2} = \frac{y^2}{4r^2}, \qquad y \in [0, 2r]$$

- We'll take the derivative of the CDF to get the following:

$$f_X(x) = \frac{\mathrm{d}\mathbb{P}(X \le x)}{\mathrm{d}x} = \frac{2x}{r^2}, \qquad x \in [0, r]$$

- Using the same approach as the previous part:

$$f_Y(y) = \frac{\mathrm{d}\mathbb{P}(Y \le y)}{\mathrm{d}y} = \frac{y}{2r^2}, \qquad y \in [0, 2r]$$

(b) We'll condition on Alex's outcome and then integrate over all the possibilities to get the marginal $\mathbb{P}(X \le Y)$ as following:

$$\begin{aligned}\mathbb{P}(X \le Y) &= \int_0^{2r} \mathbb{P}(X \le Y \mid Y = y) f_Y(y)\,\mathrm{d}y = \int_0^r \frac{y^2}{r^2} \times \frac{y}{2r^2}\,\mathrm{d}y + \int_r^{2r} 1 \times \frac{y}{2r^2}\,\mathrm{d}y \\ &= \frac{r^4 - 0}{8r^4} + \frac{4r^2 - r^2}{4r^2} = \frac{1}{8} + \frac{3}{4} = \frac{7}{8}\end{aligned}$$

Note the range within which $\mathbb{P}(X \le Y) = 1$. This allowed us to separate the integral to simplify our solution. Using this, we can get $\mathbb{P}(Y \le X)$ by the following:

$$\mathbb{P}(Y \le X) = 1 - \mathbb{P}(X \le Y) = \frac{1}{8}$$

A similar approach to the integral above could be used to verify this result.

$$\mathbb{P}(Y \le X) = \int_0^r \mathbb{P}(Y \le X \mid X = x) f_X(x)\,\mathrm{d}x = \int_0^r \frac{x^2}{4r^2} \frac{2x}{r^2}\,\mathrm{d}x = \frac{1}{2r^4} \int_0^r x^3\,\mathrm{d}x = \frac{r^4}{8r^4} = \frac{1}{8}$$

(c) Getting the CDF of U relies on the insight that for the minimum of two random variables to be greater than a value, they both need to be greater than that value. Taking the complement of this will give us the CDF of U. This allows us to get the following result. For $u \in [0,r]$:

$$\begin{aligned}\mathbb{P}(U \le u) &= 1 - \mathbb{P}(U \ge u) = 1 - \mathbb{P}(X \ge u)\mathbb{P}(Y \ge u) = 1 - \big(1 - \mathbb{P}(X \le u)\big)\big(1 - \mathbb{P}(Y \le u)\big) \\ &= 1 - \left(1 - \frac{u^2}{r^2}\right)\left(1 - \frac{u^2}{4r^2}\right) = \frac{5u^2}{4r^2} - \frac{u^4}{4r^4}\end{aligned}$$

For $u > r$, we get $\mathbb{P}(X > u) = 0$, this makes $\mathbb{P}(U \le u) = 1$.

(d) Getting the CDF of V also relies on a similar insight that for the maximum of two random variables to be smaller than a value, they both need to be smaller than that value. Using this we can get the following result for $v \in [0,r]$:

$$\mathbb{P}(V \le v) = \mathbb{P}(X \le v)\mathbb{P}(Y \le v) = \left(\frac{v^2}{r^2}\right)\left(\frac{v^2}{4r^2}\right) = \frac{v^4}{4r^4}$$

For $v \in [r, 2r]$ we have $\mathbb{P}(X \le v) = 1$, this makes

$$\mathbb{P}(V \le v) = \mathbb{P}(Y \le v) = \frac{v^2}{4r^2}.$$

For $v > 2r$ we have $\mathbb{P}(V \le v) = 1$ since CDFs of both X and Y are 1 in this range.

(e) We can subtract U from V to get this difference. Using the tail-sum formula to calculate the expectation, we can get the following result:

$$\begin{aligned}\mathbb{E}[|X - Y|] &= \mathbb{E}[V - U] = \mathbb{E}[V] - \mathbb{E}[U] = \int_0^{2r} \mathbb{P}(V > v)\,\mathrm{d}v - \int_0^{r} \mathbb{P}(U > u)\,\mathrm{d}u \\ &= \int_0^{r}\left(1 - \frac{v^4}{4r^4}\right)\mathrm{d}v + \int_r^{2r}\left(1 - \frac{v^2}{4r^2}\right)\mathrm{d}v - \int_0^{r}\left(1 - \frac{5u^2}{4r^2} + \frac{u^4}{4r^4}\right)\mathrm{d}u \\ &= \frac{19r}{20} + \frac{5r}{12} - \frac{19r}{30} = \frac{11r}{15}\end{aligned}$$

Alternatively, you could derive the density of U and V and use those to calculate the expectation. For $v \in [0,r]$:

$$f_V(v) = \frac{\mathrm{d}\mathbb{P}(V \le v)}{\mathrm{d}v} = \frac{v^3}{r^4}$$

For $v \in [r, 2r]$:

$$f_V(v) = \frac{\mathrm{d}\mathbb{P}(V \le v)}{\mathrm{d}v} = \frac{v}{2r^2}$$

Using this we can calculate $\mathbb{E}[V]$ as:

$$\mathbb{E}[V] = \int_0^{2r} v f_V(v)\,\mathrm{d}v = \frac{1}{r^4}\int_0^{r} v^4\,\mathrm{d}v + \frac{1}{2r^2}\int_r^{2r} v^2\,\mathrm{d}v = \frac{r^5}{5r^4} + \frac{8r^3 - r^3}{6r^2} = \frac{r}{5} + \frac{7r}{6r} = \frac{41r}{30}$$

To calculate $\mathbb{E}[U]$ we will use the following PDF for $u \in [0,r]$:

$$f_U(u) = \frac{\mathrm{d}\mathbb{P}(U \le u)}{\mathrm{d}u} = \frac{5u}{2r^2} - \frac{u^3}{r^4}$$

We can get the $\mathbb{E}[U]$ by the following:

$$\mathbb{E}[U] = \int_0^r u f_U(u)\,\mathrm{d}u = \int_0^r \left(\frac{5u^2}{2r^2} - \frac{u^4}{r^4}\right)\mathrm{d}u = \frac{5r^3}{6r^2} - \frac{r^5}{5r^4} = \frac{5r}{6} - \frac{r}{5} = \frac{19r}{30}$$

Combining the two results gives us the same result as above:

$$\mathbb{E}[|X-Y|] = \mathbb{E}[V-U] = \mathbb{E}[V] - \mathbb{E}[U] = \frac{41r}{30} - \frac{19r}{30} = \frac{11r}{15}$$

2 Variance of the Minimum of Uniform Random Variables

Let n be a positive integer and let $X_1,\ldots,X_n \overset{\text{i.i.d.}}{\sim} \text{Uniform}[0,1]$. Find $\operatorname{var} Y$, where

$$Y := \min\{X_1,\ldots,X_n\}.$$

Solution:

We know that the density of Y is $f(y) = n(1-y)^{n-1}$, for $y \in [0,1]$, and $\mathbb{E}[Y] = (n+1)^{-1}$. It remains to compute (via integration by parts)

$$\begin{aligned}\mathbb{E}[Y^2] &= \int_0^1 y^2 \cdot n(1-y)^{n-1}\,\mathrm{d}y = n\int_0^1 y^2(1-y)^{n-1}\,\mathrm{d}y = -y^2(1-y)^n\Big|_0^1 + 2\int_0^1 y(1-y)^n\,\mathrm{d}y \\ &= \frac{2}{n+1}\int_0^1 y(n+1)(1-y)^n\,\mathrm{d}y.\end{aligned}$$

Since $g(y) := (n+1)(1-y)^n$ is the density of the minimum of $n+1$ i.i.d. Uniform$[0,1]$ random variables, we recognize the last integral as the expectation of this minimum, which is $1/(n+2)$. Thus,

$$\mathbb{E}[Y^2] = \frac{2}{(n+1)(n+2)}$$

and so

$$\operatorname{var} Y = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \frac{2}{(n+1)(n+2)} - \frac{1}{(n+1)^2} = \frac{2(n+1)-(n+2)}{(n+1)^2(n+2)} = \frac{n}{(n+1)^2(n+2)}.$$

Fun Fact: For a non-negative random variable X with density f_X, one can extend the tail sum formula to give

$$\mathbb{E}[X^2] = \int_0^\infty x^2 f_X(x)\,\mathrm{d}x = \int_0^\infty \left(\int_0^x 2s\,\mathrm{d}s\right) f_X(x)\,\mathrm{d}x = \int_0^\infty 2s\int_s^\infty f_X(x)\,\mathrm{d}x\,\mathrm{d}s = \int_0^\infty 2s\,\mathbb{P}(X \geq s)\,\mathrm{d}s$$

and this gives another way to compute $\mathbb{E}[Y^2]$ in this problem. You can derive a similar formula to compute any moment $\mathbb{E}[X^k]$ for $k \in \mathbb{N}$.

3 Exponential Practice

Let $X \sim \text{Exponential}(\lambda_X)$ and $Y \sim \text{Exponential}(\lambda_Y)$ be independent, where $\lambda_X, \lambda_Y > 0$. Let $U = \min\{X,Y\}$, $V = \max\{X,Y\}$, and $W = V - U$.

(a) Compute $\mathbb{P}(U > t, X \le Y)$, for $t \ge 0$.

(b) Use the previous part to compute $\mathbb{P}(X \le Y)$. Conclude that the events $\{U > t\}$ and $\{X \le Y\}$ are independent.

(c) Compute $\mathbb{P}(W > t \mid X \le Y)$.

(d) Use the previous part to compute $\mathbb{P}(W > t)$.

(e) Calculate $\mathbb{P}(U > u, W > w)$, for $w > u > 0$. Conclude that U and W are independent. [*Hint*: Think about the approach you used for the previous parts.]

Solution:

(a) One has

$$\begin{aligned}
\mathbb{P}(U > t, X \le Y) = \mathbb{P}(t < X \le Y) &= \int_t^\infty \int_x^\infty f_{X,Y}(x,y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_t^\infty \int_x^\infty \lambda_X \exp(-\lambda_X x)\lambda_Y \exp(-\lambda_Y y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \lambda_X \lambda_Y \int_t^\infty \exp(-\lambda_X x) \cdot \frac{\exp(-\lambda_Y x)}{\lambda_Y}\,\mathrm{d}x = \lambda_X \int_t^\infty \exp\big(-(\lambda_X + \lambda_Y)x\big)\,\mathrm{d}x \\
&= \frac{\lambda_X}{\lambda_X + \lambda_Y} \exp\big(-(\lambda_X + \lambda_Y)t\big).
\end{aligned}$$

(b) Take $t = 0$.

$$\mathbb{P}(X \le Y) = \frac{\lambda_X}{\lambda_X + \lambda_Y}.$$

Since X and Y are independent exponentials, $U = \min\{X,Y\} \sim \text{Exponential}(\lambda_X + \lambda_Y)$. So, $\mathbb{P}(U > t) = \exp(-(\lambda_X + \lambda_Y)t)$, and therefore we have $\mathbb{P}(U > t, X \le Y) = \mathbb{P}(X \le Y)\mathbb{P}(U > t)$.

(c) One has

$$\begin{aligned}
\mathbb{P}(W > t, X \le Y) = \mathbb{P}(Y - X > t) &= \int_0^\infty \int_{x+t}^\infty \lambda_X \exp(-\lambda_X x)\lambda_Y \exp(-\lambda_Y y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \lambda_X \lambda_Y \int_0^\infty \exp(-\lambda_X x) \cdot \frac{\exp(-\lambda_Y (x+t))}{\lambda_Y}\,\mathrm{d}x \\
&= \lambda_X \exp(-\lambda_Y t) \int_0^\infty \exp\big(-(\lambda_X + \lambda_Y)x\big)\,\mathrm{d}x = \frac{\lambda_X}{\lambda_X + \lambda_Y} \exp(-\lambda_Y t).
\end{aligned}$$

So, we see that

$$\mathbb{P}(W > t \mid X \le Y) = \frac{\mathbb{P}(W > t, X \le Y)}{\mathbb{P}(X \le Y)} = \exp(-\lambda_Y t).$$

Alternatively,

$$\begin{aligned}\mathbb{P}(W > t \mid X \le Y) = \mathbb{P}(Y > X + t \mid X \le Y) &= \int_0^\infty \mathbb{P}(Y > x + t \mid Y \ge x) f_X(x)\,\mathrm{d}x \\ &= \exp(-\lambda_Y t)\int_0^\infty f_X(x)\,\mathrm{d}x = \exp(-\lambda_Y t),\end{aligned}$$

where we have used the memoryless property of the exponential distribution. Note that in the first line, we are using conditioning:

$$\mathbb{P}(Y > X + t \mid X \le Y) = \int_0^\infty \mathbb{P}(Y > X + t \mid X \le Y, X = x) f_X(x)\,\mathrm{d}x.$$

The probability inside the integral then becomes $\mathbb{P}(Y > x + t \mid Y \ge x, X = x)$, and then one can drop the conditioning on $X = x$ because X and Y are independent.

(d) By switching X and Y in the previous part, we have

$$\mathbb{P}(W > t \mid Y \le X) = \exp(-\lambda_X t).$$

So, we can use the law of total probability to give

$$\begin{aligned}\mathbb{P}(W > t) &= \mathbb{P}(X \le Y)\mathbb{P}(W > t \mid X \le Y) + \mathbb{P}(Y \le X)\mathbb{P}(W > t \mid Y \le X) \\ &= \frac{\lambda_X}{\lambda_X + \lambda_Y}\exp(-\lambda_Y t) + \frac{\lambda_Y}{\lambda_X + \lambda_Y}\exp(-\lambda_X t).\end{aligned}$$

(e) We calculate

$$\begin{aligned}\mathbb{P}(U > u, W > w, X \le Y) &= \mathbb{P}(u < X \le X + w < Y) \\ &= \int_u^\infty \int_{x+w}^\infty \lambda_X \exp(-\lambda_X x)\lambda_Y \exp(-\lambda_Y y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \lambda_X \lambda_Y \int_u^\infty \exp(-\lambda_X x)\cdot \frac{\exp(-\lambda_Y(x+w))}{\lambda_Y}\,\mathrm{d}x \\ &= \lambda_X \exp(-\lambda_Y w)\int_u^\infty \exp\bigl(-(\lambda_X + \lambda_Y)x\bigr)\,\mathrm{d}x \\ &= \frac{\lambda_X}{\lambda_X + \lambda_Y}\exp(-\lambda_Y w)\exp\bigl(-(\lambda_X + \lambda_Y)u\bigr).\end{aligned}$$

By switching the roles of X and Y in the above computation, we obtain

$$\mathbb{P}(U > u, W > w, Y \le X) = \frac{\lambda_Y}{\lambda_X + \lambda_Y}\exp(-\lambda_X w)\exp\bigl(-(\lambda_X + \lambda_Y)u\bigr).$$

Now, we add the two expressions together to obtain

$$\begin{aligned}\mathbb{P}(U > u, W > w) &= \Bigl(\frac{\lambda_X}{\lambda_X + \lambda_Y}\exp(-\lambda_Y w) + \frac{\lambda_Y}{\lambda_X + \lambda_Y}\exp(-\lambda_X w)\Bigr)\exp\bigl(-(\lambda_X + \lambda_Y)u\bigr) \\ &= \mathbb{P}(W > w)\mathbb{P}(U > u).\end{aligned}$$

So, U and W are independent!

4 Exponential Practice II

(a) Let $X_1, X_2 \sim \text{Exponential}(\lambda)$ be independent, $\lambda > 0$. Calculate the density of $Y := X_1 + X_2$. [*Hint*: One way to approach this problem would be to compute the CDF of Y and then differentiate the CDF.]

(b) Let $t > 0$. What is the density of X_1, conditioned on $X_1 + X_2 = t$? [*Hint*: Once again, it may be helpful to consider the CDF $\mathbb{P}(X_1 \le x \mid X_1 + X_2 = t)$. To tackle the conditioning part, try conditioning instead on the event $\{X_1 + X_2 \in [t, t+\varepsilon]\}$, where $\varepsilon > 0$ is small.]

Solution:

(a) Let $y > 0$. Observe that if $X_1 + X_2 \le y$, then since $X_1, X_2 \ge 0$, it follows that $X_1 \le y$ and $X_2 \le y - X_1$.

$$\begin{aligned}
\mathbb{P}(Y \le y) = \mathbb{P}(X_1 \le y, X_2 \le y - X_1) &= \int_0^y \int_0^{y-x_1} \lambda \exp(-\lambda x_1) \lambda \exp(-\lambda x_2) \, \mathrm{d}x_2 \, \mathrm{d}x_1 \\
&= \lambda^2 \int_0^y \exp(-\lambda x_1) \cdot \frac{1 - \exp(-\lambda (y - x_1))}{\lambda} \, \mathrm{d}x_1 \\
&= \lambda \int_0^y (\exp(-\lambda x_1) - \exp(-\lambda y)) \, \mathrm{d}x_1 = \lambda \left(\frac{1 - \exp(-\lambda y)}{\lambda} - y \exp(-\lambda y) \right).
\end{aligned}$$

Upon differentiating the CDF, we have

$$\begin{aligned}
f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} \mathbb{P}(Y \le y) &= \lambda \exp(-\lambda y) - \lambda \exp(-\lambda y) + \lambda^2 y \exp(-\lambda y) \\
&= \lambda^2 y \exp(-\lambda y), \qquad y > 0.
\end{aligned}$$

(b) Let $0 \le x \le t$. Following the hint, we have

$$\begin{aligned}
\mathbb{P}(X_1 \le x \mid X_1 + X_2 \in [t, t+\varepsilon]) &= \frac{\mathbb{P}(X_1 \le x, X_1 + X_2 \in [t, t+\varepsilon])}{\mathbb{P}(X_1 + X_2 \in [t, t+\varepsilon])} \\
&= \frac{\mathbb{P}(X_1 \le x, X_2 \in [t - X_1, t - X_1 + \varepsilon])}{f_Y(t) \cdot \varepsilon} \\
&= \frac{\int_0^x \int_{t-x_1}^{t-x_1+\varepsilon} \lambda \exp(-\lambda x_1) \lambda \exp(-\lambda x_2) \, \mathrm{d}x_2 \, \mathrm{d}x_1}{\lambda^2 t \exp(-\lambda t) \cdot \varepsilon} \\
&= \frac{\lambda^2 \int_0^x \exp(-\lambda x_1) \exp(-\lambda (t - x_1)) \varepsilon \, \mathrm{d}x_1}{\lambda^2 t \exp(-\lambda t) \cdot \varepsilon} = \frac{\int_0^x \mathrm{d}x_1}{t} = \frac{x}{t}.
\end{aligned}$$

This means that the density is

$$f_{X_1 \mid X_1 + X_2}(x \mid t) = \frac{\mathrm{d}}{\mathrm{d}x} \mathbb{P}(X \le x \mid X_1 + X_2 = t) = \frac{1}{t}, \qquad x \in [0, t],$$

which means that conditioned on $X_1 + X_2 = t$, X_1 is actually uniform on the interval $[0, t]$!

5 Moments of the Exponential Distribution

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

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 $t \in [0, \lambda)$,

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

and also

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

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

6 Exponential Approximation to the Geometric Distribution

Say you want to buy your friend a gift for her birthday but it totally slipped your mind and your friend's birthday already passed. Well, better late than never, so you order a package from Amazon which will arrive in X seconds, where $X \sim \text{Geometric}(p)$ (for $p \in (0,1)$). The later the package arrives, the worse it is, so say that the cost of giving your friend the gift is X^4, and you wish to compute $\mathbb{E}[X^4]$. Unfortunately, computing $\mathbb{E}[X^4]$ is very tedious for the geometric distribution, so approximate X by a suitable exponential distribution and compute $\mathbb{E}[X^4]$.

To get started on this problem, take a look at $\mathbb{P}(X > x)$ for the geometric distribution and the exponential distribution and note the similarity.

[*Note*: This problem illustrates that sometimes, moving to the continuous world simplifies calculations!]

Solution:

For each positive integer k, $\mathbb{P}(X > k) = (1-p)^k = \exp(k\ln(1-p))$, so by replacing k with a continuous variable $t > 0$, we approximate X as $X \approx \text{Exponential}(\lambda)$, where $\lambda := -\ln(1-p)$. So,

$$\mathbb{E}[X^4] \approx \frac{4!}{\lambda^4} = \frac{24}{\ln(1-p)^4}.$$

How good is this approximation? Approximating a geometric distribution via an exponential distribution is good when the time scale is small enough so that you do not notice the "discrete" aspect of the geometric distribution, i.e., when you discretize time into a large number of intervals in which the probability of success in each interval is small. Notice that the problem specifies that X is the number of *seconds* until the package arrives.