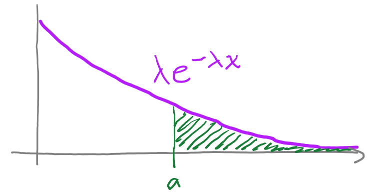
From last time:

6. Let $X \sim \text{Exp}(\lambda)$ and $0 < a < b$.



(a) What is $P(X \geq a)$?

First, the cdf:

$$P(X \leq a) = \int_0^a \lambda e^{-\lambda x}\,dx = -e^{-\lambda x}\Big|_0^a = -e^{-\lambda a} + e^0 = 1 - e^{-\lambda a}$$

Thus, $P(X \geq a) = 1 - (1 - e^{-\lambda a}) = \boxed{e^{-\lambda a}}$ ← exponential tail probability

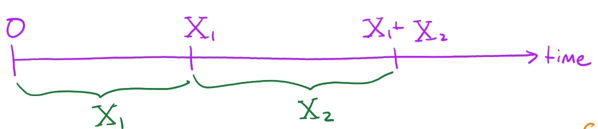
(b) Show that $\underline{P(X > b \mid X > a) = P(X > b - a)}$. ← Memoryless Property!

$$P(X > b \mid X > a) = \frac{P(X > b \text{ and } X > a)}{P(X > a)} = \frac{P(X > b)}{P(X > a)} = \frac{e^{-\lambda b}}{e^{-\lambda a}}$$

$$= e^{-\lambda(b-a)} = P(X > b - a)$$

New problems for today:

1. Suppose that the time between goals in a hockey game is exponentially distributed with mean 18 minutes (ignore timeouts and stoppages). Let X be the time from the start of the game until the second goal occurs.



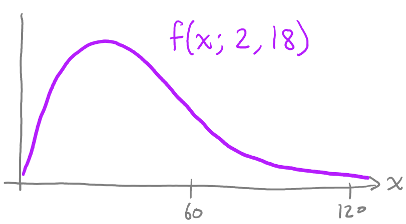
X_1 and X_2 are each $\text{Exp}\left(\frac{1}{18}\right)$ $\lambda = \frac{1}{18}$

[PREVIEW OF CH. 4: independent random variables]

$\alpha = 2,\ \beta = \frac{1}{\lambda} = 18$

Let $X = X_1 + X_2$. Then $X \sim \text{Gamma}(2, 18)$.

(a) Sketch the pdf of X.



(b) What is the probability that the second goal occurs less than 30 minutes after the game starts?

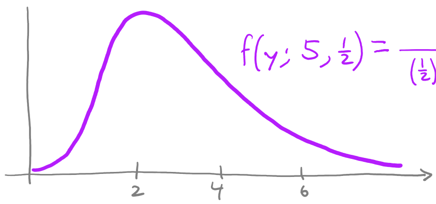
Then $P(X < 30) = \text{pgamma}\left(30, 2, \frac{1}{18}\right) = 0.496$

2. Suppose that a call center receives calls according to a Poisson distribution at a rate of 2 calls per minute. Let Y be the time between 10:00am and the 5th call received after 10:00am.

Time between calls is Exp(2).

$Y \sim \text{Gamma}(5, \frac{1}{2})$ is the sum of five Exp(2) rvs.

(a) Sketch the pdf of Y.



$f(y; 5, \frac{1}{2}) = \frac{1}{(\frac{1}{2})^5 \Gamma(5)} y^4 e^{-2y} = \frac{32}{24} y^4 e^{-2y} = \frac{4}{3} y^4 e^{-2y}$

(b) What are the mean and variance of Y?

$E(Y) = 5 \cdot \frac{1}{2} = \frac{5}{2}$

$Var(Y) = 5\left(\frac{1}{2}\right)^2 = \frac{5}{4}$

(c) What is $P(Y < 1)$?

R

$2 = \frac{1}{\beta}$

$\int_0^1 \frac{4}{3} y^4 e^{-2y}\, dy = \text{pgamma}(1, 5, 2) = \underline{0.0526}$

3. For large α, the gamma distribution converges to a normal distribution with mean $\alpha\beta$ and variance $\alpha\beta^2$. Investigate this in the case that $\beta = 1$.

(a) Let $X \sim \text{Gamma}(10, 1)$. Use technology to compute $P(X \le x)$ for various values of x.

```
Table[CDF[GammaDistribution[10, 1], x], {x, 4, 16, 2}] // N
{0.00813224, 0.083924, 0.283376, 0.54207, 0.757608, 0.890601, 0.956702}
```

(b) Let $Z \sim N(10, \sqrt{10})$. Use technology to compute $P(Z \le x)$ for the same values of x that you used in part (a). Do you find the probabilities to be close to what you found in part (a)?

```
Table[CDF[NormalDistribution[10, Sqrt[10]], x], {x, 4, 16, 2}] // N
{0.0288898, 0.102952, 0.263545, 0.5, 0.736455, 0.897048, 0.97111}
```

These probabilities are somewhat close to those in part (a).

(c) Now choose a larger value of α, such as $\alpha = 100$. Compute some probabilities to verify that $X \sim \text{Gamma}(\alpha, 1)$ has nearly the same distribution as $Z \sim N(\alpha, \sqrt{\alpha})$.

```
g = Table[CDF[GammaDistribution[100, 1], x], {x, 80, 140, 10}] // N
{0.0171083, 0.158221, 0.513299, 0.841721, 0.972136, 0.99725, 0.999839}

n = Table[CDF[NormalDistribution[100, 10], x], {x, 80, 140, 10}] // N
{0.0227501, 0.158655, 0.5, 0.841345, 0.97725, 0.99865, 0.999968}

g / n
{0.752009, 0.997263, 1.0266, 1.00045, 0.994767, 0.998598, 0.999871}
```

4. The skewness coefficient of the distribution of random variable X is defined

$$\gamma = \frac{E[(X-\mu)^3]}{\sigma^3}$$

How could you compute the skewness of $X \sim \text{Gamma}(\alpha, \beta)$? Then compute the skewness of X.

option 1: integrate $\displaystyle\int_0^\infty (x-\alpha\beta)^3 \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}\, dx = 2\beta^3\alpha$

so $\displaystyle\gamma = \frac{2\beta^3\alpha}{(\alpha\beta^2)^{3/2}} = \frac{2}{\sqrt{\alpha}}$

[NOTE: The skewness depends only on the shape parameter α.]

option 2: Mathematica

```
Skewness[GammaDistribution[α, β]]
```

$$\frac{2}{\sqrt{\alpha}}$$

BONUS: Show that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx$$

$$\Gamma\left(\tfrac{1}{2}\right) = \int_0^\infty x^{-\frac{1}{2}} e^{-x}\, dx = \int_0^\infty \frac{\sqrt{2}}{y} e^{-y^2/2}\, y\, dy = \sqrt{2}\int_0^\infty e^{-y^2/2}\, dy = \sqrt{2}\cdot\frac{\sqrt{\pi}}{\sqrt{2}} = \sqrt{\pi}$$

substitute: $y = \sqrt{2x}$

$x = \frac{y^2}{2}$

$dx = y\, dy$

standard normal pdf:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-x^2/2}\, dx = 1$$

so $\displaystyle\int_0^\infty e^{-x^2/2}\, dx = \frac{\sqrt{2\pi}}{2} = \frac{\sqrt{\pi}}{\sqrt{2}}$

1. An interviewer is given a long list of people that she can interview. When asked, suppose that each person independently agrees to be interviewed with probability 0.45. The interviewer must conduct ten interviews. Let X be the number of people she must ask to be interviewed in order to obtain ten interviews.

   (a) What is the probability that the interviewer will obtain ten interviews by asking no more than 18 people?

   X is negative binomial rv with $p = 0.45$ and $r = 10$.

   $P(X \le 18) = \text{pnbinom}(8, 10, 0.45) = 0.2527$

   (b) What are the expected value and variance of the number of people who *decline* to be interviewed before the interviewer finds ten who agree?

   $E(X-10) = E(X) - 10 = \frac{10}{0.45} - 10 = 12.22$

   $Var(X-10) = Var(X) = \frac{10(1-0.45)}{(0.45)^2} = 27.16$

2. Let $X \sim \text{Geom}(p)$. Find the expected value of $\frac{1}{X}$.

   X has mass function $p(x) = (1-p)^{x-1}p$ for $x = 1, 2, 3, \ldots$

   Then $E\left(\frac{1}{X}\right) = \underbrace{\sum_{x=1}^{\infty} \frac{1}{x}(1-p)^{x-1}p}_{} = \underbrace{\frac{p\ln(p)}{p-1}}_{}$

   To see this, start with the geometric series $\sum_{x=0}^{\infty} r^x = \frac{1}{1-r}$.

   Integrate both sides, and do some algebra.

   Or, use Mathematica or Wolfram Alpha to evaluate the sum.

3. Suppose that $X \sim \text{Exp}(3)$, and let $Y = \lfloor X \rfloor$ denote the largest integer that is less than or equal to X. For example, $\lfloor 2.1 \rfloor = 2$, $\lfloor 5.99 \rfloor = 5$, and $\lfloor 14 \rfloor = 14$.

   (a) Is Y a discrete or continuous random variable?

   Possible values of Y are $0, 1, 2, 3, \ldots$, so Y is discrete.

   (b) Find $P(Y \le 1)$.

   $P(Y \le 1) = P(X < 2) = \int_0^2 3e^{-3x}\,dx = -e^{-3x}\Big|_0^2 = 1 - e^{-6} \approx 0.9975$

   (c) Find $P(Y = 2)$.

   $P(Y = 2) = P(2 \le X < 3) = \int_2^3 3e^{-3x}\,dx = -e^{-3x}\Big|_2^3 = e^{-6} - e^{-9} \approx 0.0023$

(d) Can you generalize? What is $P(Y = n)$, for any positive integer n? Is the distribution of Y one of the distributions that we have studied in this course?

$$P(Y=n) = P(n \le X < n+1) = \int_n^{n+1} 3e^{-3x}\,dx = -e^{-3x}\Big|_n^{n+1} = -e^{-3(n+1)} + e^{-3n} = e^{-3n}(1-e^{-3}) = (1-p)^n p,$$

where $p = 1 - e^{-3}$.

This is almost the pmf of a geometric random variable.

In fact, $Y+1$ has a geometric distribution with $p = 1-e^{-3}$.

4. Let $X \sim \text{Unif}[0,1]$. Compute the nth moment of X in two different ways.

(a) Use the formula $E(X^n) = \int_0^1 x^n dx$.

$$E(X^n) = \int_0^1 x^n\,dx = \frac{x^{n+1}}{n+1}\Big|_0^1 = \frac{1}{n+1}$$

(b) Use the moment generating function $M_X(t)$.

$$M_X(t) = \begin{cases} \dfrac{e^t - 1}{t} & \text{if } t \ne 0 \\ 1 & \text{if } t = 0 \end{cases}$$

Recall that $e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!} = 1 + t + \frac{t^2}{2} + \frac{t^3}{3!} + \cdots$.

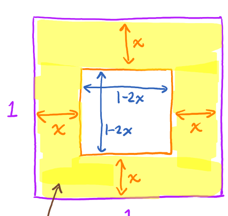
Thus, as a power series, $M_X(t) = \sum_{n=1}^{\infty} \frac{t^{n-1}}{n!}$.

Reindexing, $M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{(n+1)!} = \sum_{n=0}^{\infty} \frac{1}{n+1} \cdot \frac{t^n}{n!}$.

The coeffient of $\frac{t^n}{n!}$ in the power series is $E(X^n)$, so $E(X^n) = \frac{1}{n+1}$.

5. Choose a point uniformly at random in a unit square (i.e., a square of side length 1.) Let X be the distance from the point chosen to the nearest edge of the square. Find the cdf of X. (*Hint*: draw a picture!) Then find the pdf of X.

First, find the cdf of X. It's helpful to draw a picture:



$P(X \le x)$ is this area.

If $x \in [0, \frac{1}{2})$, then:

$$F_X(x) = P(X \le x) = 1 - (1-2x)^2 = 4x - 4x^2$$

Thus, the pdf is

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx}(4x - 4x^2) = 4 - 8x \quad \text{for } 0 \le x \le \tfrac{1}{2}.$$