

Name (please print) NASSER ABBASI

1. Suppose that two components have independent exponentially distributed lifetimes T_1 and T_2 , with parameters α and β , respectively. Find (a) $P(T_1 > T_2)$, (b) identify the distribution of $W = 2T_2$, and (c) use the results in parts (a) and (b) to obtain $P(T_1 > 2T_2)$.

(a) Problem review:

T_1 is a random variable and T_2 is a random variable, where $T_1 \sim \alpha e^{-\alpha t_1}$ and $T_2 \sim \beta e^{-\beta t_2}$.

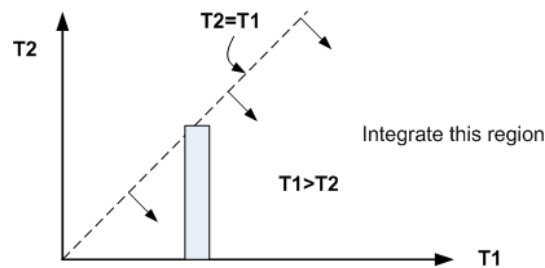
α and β can be thought of as the failure rate for each respective component. T_i is the lifetime of component i . Hence $P(T_1 = t_1)$ means to ask for the probability of the first component to have a lifetime of t_1 given that the failure rate of this kind of components is α .

solution:

Now we know that

$$P(T_1 > T_2) = \int \int f_{T_1, T_2}(t_1, t_2) dt_2 dt_1$$

Looking at the following diagram to help determine the region to integrate:



Hence

$$P(T_1 > T_2) = \int_{t_1=0}^{t_1=\infty} \int_{t_2=0}^{t_2=t_1} f_{T_1, T_2}(t_1, t_2) dt_2 dt_1$$

But since $T_1 \perp T_2$, then the joint density is the product of the marginal densities.

Hence

$$\begin{aligned} f_{T_1, T_2}(t_1, t_2) &= f_{T_1}(t_1) f_{T_2}(t_2) \\ &= \alpha e^{-\alpha t_1} \beta e^{-\beta t_2} \end{aligned}$$

Therefore

$$\begin{aligned}
P(T_1 > T_2) &= \int_0^\infty \int_0^{t_1} \alpha e^{-\alpha t_1} \beta e^{-\beta t_2} dt_2 dt_1 \\
&= \beta \alpha \int_0^\infty e^{-\alpha t_1} \left(\int_0^{t_1} e^{-\beta t_2} dt_2 \right) dt_1 \\
&= \beta \alpha \int_0^\infty e^{-\alpha t_1} \left(-\frac{1}{\beta} \left[e^{-\beta t_2} \right]_{t_2=0}^{t_2=t_1} \right) dt_1 \\
&= -\alpha \int_0^\infty e^{-\alpha t_1} \left[e^{-\beta t_1} - 1 \right] dt_1 \\
&= -\alpha \int_0^\infty e^{t_1(-\alpha-\beta)} - e^{-\alpha t_1} dt_1 \\
&= -\alpha \left(\left[\frac{1}{(-\alpha-\beta)} e^{t_1(-\alpha-\beta)} \right]_0^\infty - \frac{1}{-\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right)
\end{aligned}$$

We take $\alpha, \beta \geq 0$ since we expect the lifetime to go to zero eventually. Also this is a requirement for the integrals to not diverge.

Hence the above becomes

$$\begin{aligned}
P(T_1 > T_2) &= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[e^{t_1(-\alpha-\beta)} \right]_0^\infty + \frac{1}{\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right) \\
&= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[e^{-\infty} - 1 \right] + \frac{1}{\alpha} \left[e^{-\infty} - 1 \right] \right) \\
&= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[0 - 1 \right] + \frac{1}{\alpha} \left[0 - 1 \right] \right) \\
&= -\alpha \left(\frac{1}{(\alpha+\beta)} - \frac{1}{\alpha} \right) \\
&= -\alpha \left(\frac{\alpha - (\alpha+\beta)}{\alpha(\alpha+\beta)} \right) \\
&= -\left(\frac{\alpha - \alpha - \beta}{(\alpha+\beta)} \right)
\end{aligned}$$

Hence

$$\boxed{P(T_1 > T_2) = \frac{\beta}{(\alpha+\beta)}}$$

(b)

$$\begin{aligned}
F_W(w) &= P(W \leq w) \\
&= P(2T_2 \leq w) \\
&= P\left(T_2 \leq \frac{w}{2}\right) \\
&= F_{T_2}\left(\frac{w}{2}\right)
\end{aligned}$$

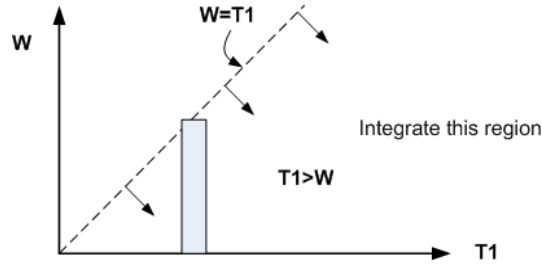
Hence

$$f_W(w) = f_{T_2}\left(\frac{w}{2}\right) \times \frac{d}{dw}\left(\frac{w}{2}\right)$$

Hence

$$f_W(w) = \frac{1}{2} f_{T_2}\left(\frac{w}{2}\right)$$

(c) Need to find $P(T_1 > 2T_2)$ which is the same as $P(T_1 > W)$, hence this is the same as part(a) but replace T_2 by W as show in the following diagram



Hence

$$\begin{aligned}
 P(T_1 > W) &= \int_0^\infty \int_0^{t_1} f_{T_1}(t_1) f_W(w) dw dt_1 \\
 &= \int_0^\infty \int_0^{t_1} f_{T_1}(t_1) \left[\frac{1}{2} f_{T_2}\left(\frac{w}{2}\right) \right] dw dt_1 \\
 &= \int_0^\infty \int_0^{t_1} \alpha e^{-\alpha t_1} \left[\frac{1}{2} \beta e^{-\beta\left(\frac{w}{2}\right)} \right] dw dt_1 \\
 &= \frac{1}{2} \beta \alpha \int_0^\infty e^{-\alpha t_1} \left(\int_0^{t_1} e^{-\beta\left(\frac{w}{2}\right)} dw \right) dt_1 \\
 &= \frac{1}{2} \beta \alpha \int_0^\infty e^{-\alpha t_1} \left(-\frac{2}{\beta} \left[e^{-\beta\left(\frac{w}{2}\right)} \right]_{w=0}^{w=t_1} \right) dt_1 \\
 &= -\alpha \int_0^\infty e^{-\alpha t_1} \left[e^{-\beta\left(\frac{t_1}{2}\right)} - 1 \right] dt_1 \\
 &= -\alpha \int_0^\infty e^{t_1\left(-\alpha - \frac{\beta}{2}\right)} - e^{-\alpha t_1} dt_1 \\
 &= -\alpha \int_0^\infty e^{t_1\left(\frac{-2\alpha - \beta}{2}\right)} - e^{-\alpha t_1} dt_1 \\
 &= -\alpha \left(\left[\frac{2}{(-2\alpha - \beta)} e^{t_1\left(\frac{-2\alpha - \beta}{2}\right)} \right]_0^\infty - \frac{1}{-\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right) \\
 &= -\alpha \left(\frac{2}{(-2\alpha - \beta)} [0 - 1] + \frac{1}{\alpha} [0 - 1] \right) \\
 &= -\alpha \left(\frac{2}{(2\alpha + \beta)} - \frac{1}{\alpha} \right)
 \end{aligned}$$

Hence

$$\begin{aligned}
 P(T_1 > W) &= - \left(\frac{2\alpha - (2\alpha + \beta)}{(2\alpha + \beta)} \right) \\
 &= - \left(\frac{2\alpha - 2\alpha - \beta}{(2\alpha + \beta)} \right)
 \end{aligned}$$

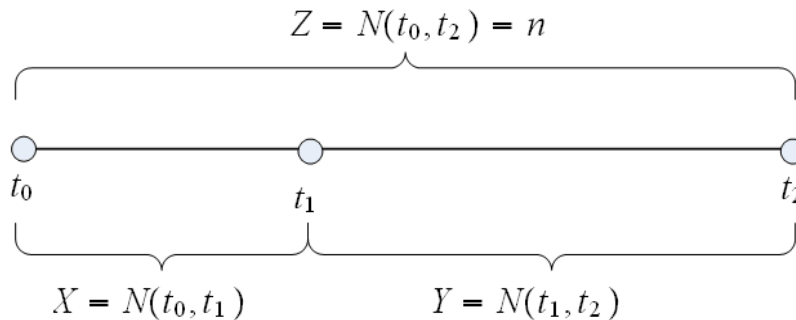
Then

$$P(T_1 > W) = \frac{\beta}{(2\alpha + \beta)}$$

2. Consider a Poisson process on the real line, and denote by $N(t_1, t_2)$ the number of events in the interval (t_1, t_2) . If $t_0 < t_1 < t_2$, find the conditional distribution of $N(t_0, t_1)$ given that $N(t_0, t_2) = n$, and identify the distribution.

Problem review: Poisson probability density is a discrete probability function (We normally call it the probability mass function *pmf*). This means the random variable is a discrete random variable.

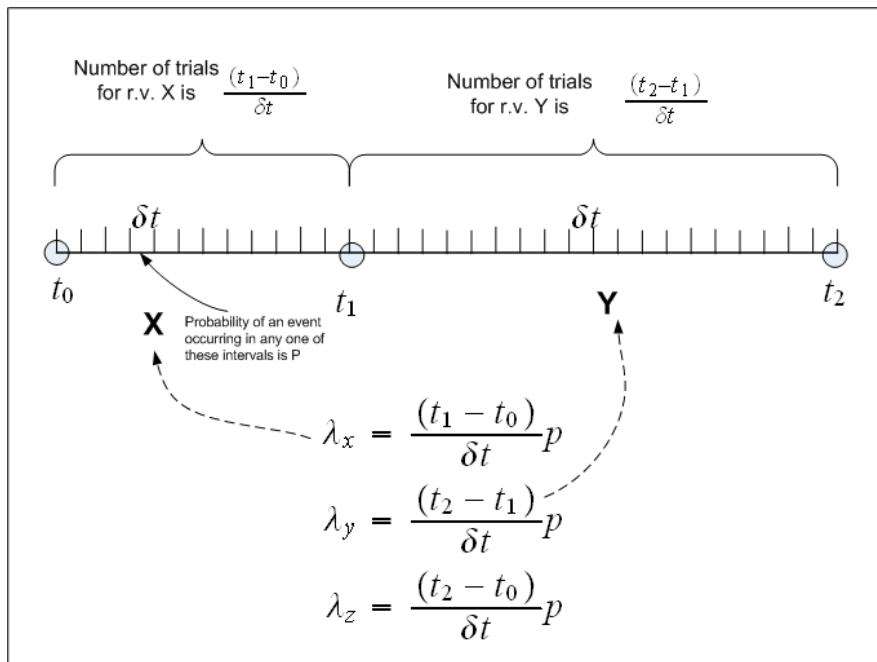
The random variable X in this case is the number of success in n trials where the probability of success in each one trial is p and the trials are independent from each others. The difference between Poisson and Binomial is that in Poisson we are looking at the problem as n becomes very large and p becomes very small in such a way that the product np goes to a fixed value which is called λ , the Poisson parameter. And then we write $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ where $k = 0, 1, 2, \dots$. The following diagram illustrates this problem, showing the three r.v. we need to analyze and the time line.



But what is "trials" in this problem? If we divide the time line itself into very small time intervals δt then the number of time intervals is the number of trials, and we assume that at most one event will occur in this time interval (since it is too small). The probability p of event occurring in this δt is the same in the interval $[t_0, t_1]$ and in the interval $[t_1, t_2]$. Now let us find λ for X and Y and Z based on this. Since $\lambda = np$ where n is the number of trials, then for X we have $\lambda_x = np = \frac{(t_1 - t_0)}{\delta t} p$ where we divided the time interval by the time

width δt to obtain the number of time slots for X . We do the same for Y and obtain that $\lambda_y = \frac{(t_2 - t_1)}{\delta t} p$

Similarly, $\lambda_Z = \frac{(t_2 - t_0)}{\delta t} p = \frac{(t_2 - t_1) + (t_1 - t_0)}{\delta t} p = \frac{(t_2 - t_1)}{\delta t} p + \frac{(t_1 - t_0)}{\delta t} p$, hence $\lambda_z = \lambda_x + \lambda_y$



Let us refer to the random variable $N(t_1, t_2)$ as Y and the r.v. $N(t_0, t_1)$ as X and the r.v. $N(t_0, t_2)$ as Z . The problem is then asking to find $P(X = x|Z = n)$ and to identify $pmf(X|Z)$. To help in the solution, we first draw a diagram to make it more clear. We take λ to be the same for the 3 random variables X, Y, Z .

$$P(X = x|Z = n) = \frac{P(X = x, Z = n)}{P(Z = n)}$$

But $Z = n$ is the same as $X + Y = n$ hence

$$\begin{aligned} P(X = x|Z = n) &= \frac{P(X = x, (X + Y) = n)}{P(Z = n)} \\ &= \frac{P(X = x, Y = n - x)}{P(Z = n)} \end{aligned}$$

Now r.v. $X \perp Y$, since the number of events in $[t_0, t_1]$ is independent from the number of events that could occur in $[t_1, t_2]$.

Given this, we can now write the joint probability of X, Y as the product of the marginal probabilities. Hence the numerator in the above can be rewritten and we obtain

$$P(X = x|Z = n) = \frac{P(X = x) P(Y = n - x)}{P(Z = n)} \quad (1)$$

Now since each of the above is a Poisson process, then

$$\begin{aligned} P(X = x) &= \frac{(\lambda_x)^x}{x!} e^{-\lambda_x} \\ P(Y = n - x) &= \frac{(\lambda_y)^{n-x}}{(n-x)!} e^{-\lambda_y} \\ P(Z = n) &= \frac{(\lambda_z)^n}{n!} e^{-\lambda_z} \end{aligned}$$

Hence (1) becomes

$$P(X = x|Z = n) = \left(\frac{(\lambda_x)^x}{x!} e^{-\lambda_x} \right) \left(\frac{(\lambda_y)^{n-x}}{(n-x)!} e^{-\lambda_y} \right) \frac{1}{\frac{(\lambda_z)^n}{n!} e^{-\lambda_z}} \quad (2)$$

Hence

$$P(X = x|Z = n) = \frac{n!}{x!(n-x)!} \left((\lambda_x)^x e^{-\lambda_x} \right) \left((\lambda_y)^{n-x} e^{-\lambda_y} \right) \frac{e^{\lambda_z}}{(\lambda_z)^n}$$

But we found that $\lambda_z = \lambda_x + \lambda_y$, hence the exponential term above vanishes and we get

$$\begin{aligned}
P(X = x|Z = n) &= \frac{n!}{x!(n-x)!} \frac{(\lambda_x)^x (\lambda_y)^{n-x}}{(\lambda_z)^n} \\
&= \binom{n}{x} \frac{(\lambda_x)^x (\lambda_y)^{n-x}}{(\lambda_z)^n} \\
&= \binom{n}{x} \frac{(\lambda_x)^x (\lambda_y)^{n-x}}{(\lambda_x + \lambda_y)^n} \\
&= \binom{n}{x} \frac{(\lambda_x)^x (\lambda_y)^{n-x}}{(\lambda_x + \lambda_y)^x (\lambda_x + \lambda_y)^{n-x}} \\
&= \binom{n}{x} \frac{(\lambda_x)^x}{(\lambda_x + \lambda_y)^x} \frac{(\lambda_y)^{n-x}}{(\lambda_x + \lambda_y)^{n-x}} \\
&= \binom{n}{x} \left(\frac{\lambda_x}{\lambda_x + \lambda_y} \right)^x \left(\frac{\lambda_y}{\lambda_x + \lambda_y} \right)^{n-x}
\end{aligned}$$

Let $k = \frac{\lambda_x}{\lambda_x + \lambda_y}$, then $1 - k = 1 - \frac{\lambda_x}{\lambda_x + \lambda_y} = \frac{\lambda_x + \lambda_y - \lambda_x}{\lambda_x + \lambda_y} = \frac{\lambda_y}{\lambda_x + \lambda_y}$ hence the last line above can be written as

$$\begin{aligned}
P(X = x|Z = n) &= \binom{n}{x} \left(\frac{\lambda_x}{\lambda_x + \lambda_y} \right)^x \left(1 - \frac{\lambda_x}{\lambda_x + \lambda_y} \right)^{n-x} \\
&= \binom{n}{x} (k)^x (1 - k)^{n-x}
\end{aligned}$$

But this is a Binomial with parameters n, k , hence

$$\boxed{P(X = x|Z = n) \sim \text{Binomial} \left(n, \frac{\lambda_x}{\lambda_x + \lambda_y} \right)}$$

3. Suppose that the probability Θ of getting heads for a coin is unknown, and let the prior opinion about Θ be represented by the uniform distribution on $[0,1]$. You spin the coin repeatedly and record the number of times N until a heads comes up. (a) Find the posterior density of Θ given N . (b) Use Matlab or any other software to plot the posterior for cases where $N = 1$, $N = 2$, and $N = 6$. Using your plots, explain what you infer about the probability of heads in each circumstance.

part (a)

Let θ , the probability of getting heads, be the specific value that the random number Θ can take.

Let $g(\theta)$ be the probability density of Θ , which we are told to be $U[0,1]$, and let $pmf_X(x)$ be the probability mass function of the random variable X where X is the number of times until a head first comes up. X is then a geometric random variable with parameter θ , hence

$$pmf_X(N) = P(X = N) = (1 - \theta)^{N-1} \theta \quad N = 1, 2, 3, \dots$$

The posterior density of Θ given N is then

$$h(\Theta = \theta | X = N) = \frac{pmf_X(N | \Theta = \theta) g(\theta)}{\int_0^1 pmf_X(N | \Theta = \theta) g(\theta) d\theta}$$

But

$$pmf_X(N | \Theta = \theta) = (1 - \theta)^{N-1} \theta$$

and $g(\theta) = 1$ since $\Theta \sim U[0, 1]$

Hence

$$h(\Theta = \theta | X = N) = \frac{(1 - \theta)^{N-1} \theta}{\int_0^1 (1 - \theta)^{N-1} \theta d\theta} \quad (1)$$

But Θ is a random continuous variable from $[0, 1]$, so how to evaluate the above? I can evaluate the above for different values of Θ on the real line from $[0, 1]$, and the more values I take between 0, 1 the more accurate $h(\Theta = \theta | X = N)$ will become.

Part(b)

First let me evaluate eq (1) for $N = 1, N = 2, N = 6$

For $N = 1$

$$h(\Theta = \theta | X = 1) = \frac{\theta}{\int_0^1 \theta d\theta} = \frac{\theta}{\left[\frac{\theta^2}{2}\right]_0^1} = \boxed{2\theta}$$

For $N = 2$

$$\begin{aligned} h(\Theta = \theta | X = 2) &= \frac{(1 - \theta) \theta}{\int_0^1 (1 - \theta) \theta d\theta} = \frac{(1 - \theta) \theta}{\int_0^1 (\theta - \theta^2) d\theta} = \frac{(1 - \theta) \theta}{\left[\frac{\theta^2}{2}\right]_0^1 - \left[\frac{\theta^3}{3}\right]_0^1} \\ &= \frac{(1 - \theta) \theta}{\frac{1}{2} - \frac{1}{3}} = \boxed{6(1 - \theta) \theta} \end{aligned}$$

For $N = 6$

$$h(\Theta = \theta | X = 6) = \frac{(1 - \theta)^{6-1} \theta}{\int_0^1 (1 - \theta)^{6-1} \theta d\theta} = \frac{(1 - \theta)^5 \theta}{\int_0^1 (1 - \theta)^5 \theta d\theta}$$

We can use integration by parts for the denominator, where $u = \theta, dv = (1 - \theta)^5$, when we do this we obtain

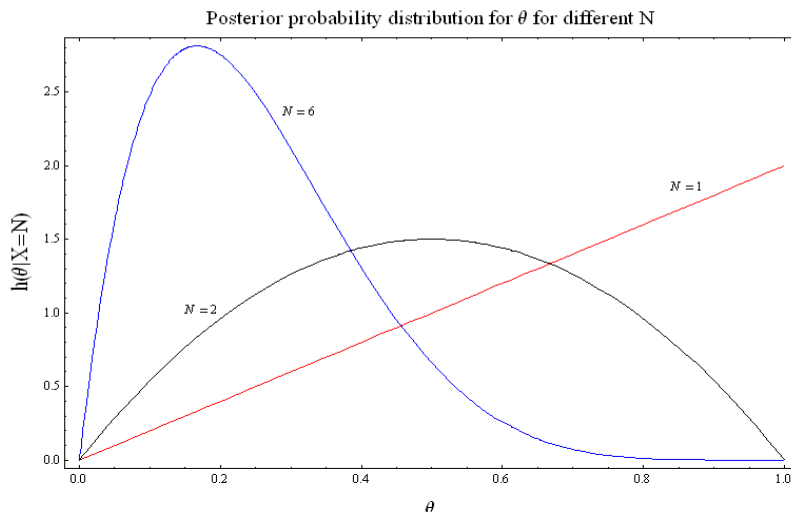
$$h(\Theta = \theta | X = 6) = 42(1 - \theta)^5 \theta$$

Now we plot the above 3 cases on the same plot:

```

In[93]:= f[n_, x_] := Which[n == 1, 2 x, n == 2, 6 (1 - x) x, n == 6, 42 (1 - x)^5 x]
Plot[{f[1, x], f[2, x], f[6, x]}, {x, 0, 1},
Frame -> True, PlotStyle -> {Red, Black, Blue},
FrameLabel -> {Style["e", 14], Style["h(θ|X=N)", 16]},
Style["Posterior probability distribution for θ for different N", 16]},
ImageSize -> 600]

```



What the above plot is saying is the following:

If it takes 'longer' to see a head comes up ($N = 6$), then the coin is taken as biased towards a tail, and the probability of getting a head becomes smaller, this is why we see that the most likely probability in this case to be around 0.15 (looking at the $N=6$ curve). We say that based on the observation of $N = 6$, then the coin has a higher probability of having its probability of getting a head to be about 0.15 than any other value. (The area around $\theta = 1.5$ is larger than any other area for the same $\delta\theta$)

Now, when $N = 2$, i.e. we flipped the coin 2 times, and got a head on the second time, then we see from the $N = 2$ curve that the coin has a most likelihood of having a probability of getting a head to be 0.5.

This is what we would expect, since in an unbiased coin, the probability of getting a head is $\frac{1}{2}$, and hence with a fair coin, we expect to see a head half of the times it is flipped, and since we flipped 2 times, and saw a head the second time, this posterior probability has its most likely value to be around .5 as well.

When $N = 1$, this says that we got a head in the first time we flipped the coin. We see that the posterior probability of getting a head now has it maximum around 1. This means the posterior probability is saying this coin is biased towards a head.

The above is a method to estimate the probability distribution of the probability itself of getting a head based on the observed events and based on the prior known probability of getting a head. Hence the events observed allow us to estimate the probability of getting a head. Hence the posterior probability is conditioned on each event as in this problem.