Lecture 13 : The Exponential Distribution

Definition

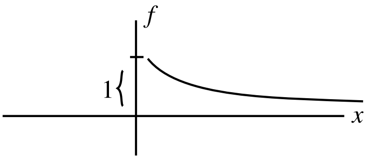
A continuous random variable X is said to have exponential distribution with parameter λ.
If the pdf of X is (with $\lambda > 0$)

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, x > 0 \\ 0, \qquad \textit{otherwise} \end{cases} \qquad (*)$$

Remarks

Very often the independent variable will be time t rather than x. The exponential distribution is the special case of the gamma distribution with $\alpha = 1$ and $\beta = \frac{1}{\alpha}$.
We will see that X β closely tied to the Poisson process, that is why λ is used above.

Here is the graph of f



Proposition ((cdf) (Prove this))

If X has exponential distribution then

$$F(x) = P(X \le x) = 1 - e^{-\lambda x}$$

Corollary (Prove this)

If X has exponential distribution

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

This is a very useful formula.

Proposition

If X has exponential distribution

(i) $E(X) = \dfrac{1}{\lambda}$

(ii) $V(X) = \dfrac{1}{\lambda^2}$

The Physical Meaning of the Exponential Distribution

Recall (Lecture 8) that the binomial process (having a a child, flipping a coin) gave rise to two (actually infinitely many) more distributions

1. X = the geometric distribution
   = the waiting time for the first girl

and X_r = the negative binomial
= the waiting time for the r-th girl

Remark

Here time was discrete. Also X_r was the number of boys before the r-th girl so the waiting time was actually $Y_r = X_r + r - 1$.

Now we will see the same thing happens with a Poisson process. Now time is continuous, as I warned you. I will switch from x to t in (*).

So suppose we have a trap to catch some species of animal. We run it forever starting at time $t = 0$, so $0 \le t < \infty$.

The Counting Random Variable

Now fix a time period t. So we have a "counting random variable X_t".
$X_t = \sharp$ of animals caught in the trap in time t.
We will choose the model $X_t \sim P(\lambda t) =$ Poisson with parameter λt.
Z We are using λ instead of α in the Poisson process
N.B. $P(X_t = 0) = e^{-\lambda t}(\sharp)$

Remark

The analogue from before was
$X_n = \sharp$ *of girls in the first n children*
(so we have a discrete "time period", the binomial random variable was the counting random variable.)
Now we want to consider the analogue of the "waiting time" random variables, the geometric and negative binomials for the binomial process.
Let Y = *the time when the first animal is caught.*

The proof of the following theorem involves such a beautiful simple idea I am going to give it.

Theorem

Y has exponential distribution with parameter α.

Proof

We will compute $P(Y > t)$ *and show* $P(Y > t) = e^{-\lambda t}$

$$(so\ \ F(t) = P(Y \leq t) = 1 - e^{-\lambda t}$$
$$and\ \ f(t) = F'(\lambda) = \lambda e^{-\lambda t})$$

Proof (Cont.)

Here is the beautiful observation. You have to wait longer than t units for the first animal to be caught

$\Leftrightarrow$ *there are no animals in the trap at time t.*

In symbols this says

$$(Y > t) = (X_t = 0)$$

equality of events

But we have seen

$$P(X_t = 0) = e^{-\lambda t} \quad \text{so necessarily}$$

$$P(Y > t) = e^{-\lambda t}$$

$\square$

Now what about the analogue of the negative binomial = the waiting time for the n-th girl.

The r-Erlang Distribution

Let $Y_r =$ the waiting until the r-th animal is caught.

Theorem

(i) *The cdf F_r of Y_r is given by*

$$F_r(t) = \begin{cases} 1 - (1 + \lambda + \frac{(\lambda t)^2}{2!} + \cdots + \frac{(\lambda t)^{r-1}}{(r-1)!}, e^{-\lambda t}, t > 0 \\ 0, \quad \textit{otherwise} \end{cases}$$

(ii) *Differentiating (this is tricky) $F_r(t)$ to get the pdf $f_r(t)$ we get*

$$f_r(t) = \begin{cases} \frac{\lambda^r t^{r-1}}{(r-1)!} e^{-\lambda t}, & t > 0 \\ 0, & \textit{otherwise} \end{cases}$$

Remark

This distribution is called the r-Erlang distribution.

Proof

We use the same trick as before

$$P(Y_r > t) = P(X_t \leq r - 1)$$

The waiting time for the r-th animal to arrive in the trap is $> t \Leftrightarrow$ *at time t there are* $\leq r - 1$ *animals in the trap.*
Since $X_t \sim P(\lambda t)$ *we have*

$$P(X_t \leq r - 1) = e^{-\lambda t} + e^{-\lambda t}\lambda t + e^{-\lambda t}\frac{(\lambda t)^2}{2!} + \cdots + e^{-\lambda t}\frac{(\lambda t)^{r-1}}{(r-1)!}$$

Proof (Cont.)

Now we have to do some hard computation.

$$P(X_t \leq r-1) = e^{-\lambda t}\left(1 + \lambda t + \cdots + \frac{(\lambda t)^{r-1}}{(r-1)!}\right)$$

So

$$\begin{aligned} F_r(t) &= P(Y_r \leq t) = 1 - P(Y_r > t) \\ &= 1 - e^{-\lambda t}\left(1 + \lambda t + \cdots + \frac{(\lambda t)^{r-1}}{(r-1)!}\right) \end{aligned}$$

But $\quad f_r(t) = \dfrac{dF_r}{dt}(t)$
So we have to differentiate the expression on the right-hand side

$$\textit{Of course} \quad \frac{d}{dt}(1) = 0$$

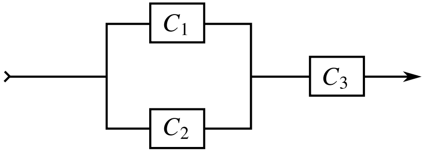
Proof (Cont.)

A hard derivative computation

$$
\begin{aligned}
f_r(t) &= -\frac{-d}{dt}(e^{-\lambda t})\left(1+\lambda t+\frac{(\lambda t)^2}{2!}+\cdots+\frac{(\lambda t)^{r-2}}{(r-2!)}+\frac{(\lambda t)^{r-?}}{(r-1)!}\right)\\
&\quad - e^{-\lambda t}\frac{d}{dt}\left(1+\lambda t+\frac{(\lambda t)^2}{1!}+\cdots+\frac{(\lambda t)^{r-1}}{(r-1)?}\right)\\
&= \lambda e^{-\lambda t}\left(1+\lambda t+\frac{(\lambda t)^2}{2!}+\cdots+\frac{(\lambda t)^{r-2}}{(r-2)!}+\frac{(\lambda t)^{r-1}}{(r-1)!}\right)\\
&\quad - e^{-\lambda t}\left(\lambda+\lambda^2 t+\frac{\lambda^3 t^2}{2!}+\cdots+\frac{\lambda^{r-1}t^{r-2}}{(r-2)!}\right)\\
&= e^{-\lambda t}\left(\cancel{\lambda}+\cancel{\lambda^2}++\cancel{\frac{\lambda^3 t^2}{2!}}+\cdots+\cancel{\frac{\lambda^{r-1}t^{r-2}}{(r-2)!}}+\frac{\lambda^r t^{r-1}}{(r-1)!}\right)\\
&\qquad\quad \updownarrow \quad \updownarrow \qquad\quad \updownarrow \qquad\qquad\qquad \updownarrow\\
&\quad - e^{-\lambda t}\left(\cancel{\lambda}+\cancel{\lambda^2}++\cancel{\frac{\lambda^2 t^2}{2!}}+\cdots+\cancel{\frac{\lambda^{r-1}t^{r-2}}{(r-2)!}}\right)\\
&= \frac{\lambda^r t^{r-1}}{(r-1)!}e^{-\lambda t}
\end{aligned}
$$

□

Lifetimes and Failure Times for Components and Systems



Suppose each of the components has a lifetime that is exponentially distributed with parameter λ (see below for a more precise statement).
Assume the components are independent. How is the system lifetime distributed?

Solution

Define random variables X_1, X_2, X_3 *by*

$$(X_i = t) = (C_i \text{ fails at time } t), i = 1, 2, 3$$

Then X_i *is exponentially distributed with parameter* λ *so*

$$P(X_i \leq t) = 1 - e^{-\lambda t}, i = 1, 2, 3$$
$$P(X_i > t) = e^{-\lambda}, i = 1, 2, 3.$$

Define Y *by*

$$(Y = t) = (\text{system fails at time } t)$$

Solution (Cont.)

The key step (using the geometry of the system)
Lump C_1 and C_2 into a single component A and let W be the corresponding random variable so

$$(W = t) = (A \ \text{ fails at time } t)$$

$$(Y > t) = (W > t) \cap (X_3 > t)$$

(the system is working at time t $\Leftrightarrow$ both A and C_3 are working at time t)

The Golden Rule

Try to get $\cap$ instead of $\cup$ - that's why I choose $(Y > t)$ on the left.
Hence

$$P(Y > t) = P((W > t) \cap (X_3 > t))$$
$$\text{by independence}$$
$$= P(W > t) \cdot P(X_3 > t) \qquad (\sharp)$$

Why are $(W > t)$ and $(X_3 > t)$ independent?

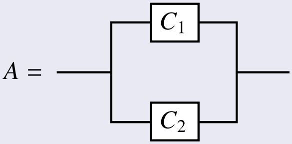
Answer

Suppose $C_1, C_2, \ldots C_n$ are independent components. Suppose
A = a subcollection of the C_i's.
B = another subcollection of the C_i's.

Answer (Cont.)

Then A and B are independent $\Leftrightarrow$ they have no common component.
So now we need

$$P(W > t) \quad \textit{where} \quad W \quad \textit{is}$$



I should switch to $P(W \leq t)$ to get intersections but I won't to show you why unions give extra terms.

$$(W > t) = (X_1 > t) \cup (X_2 > t)$$

(A is working at time $t \Leftrightarrow$ either C_1 is or C_2 is)
So

extra term

$$\begin{aligned} P(W > t) &= P(X_1 > t) + P(X_2 > t) \\ &\quad - P((X_1 > t) \cap (X_1 > t)) \\ &= P(X_1 > t) + P(X_2 > t) \\ &\quad P(X_1 > t)P(X_2 > t) \\ &= e^{-\lambda t} t e^{-\lambda t} - (e^{-\lambda t})(e^{-\lambda t}) \\ &= Z e^{-\lambda t} - e^{-2\lambda t} \end{aligned}$$

Now from (♯)

$$
\begin{aligned}
P(Y > t) &= P(W > t)P(X_3 > t) \\
&= (2e^{-\lambda t} - e^{-2\lambda t})e^{-\lambda t} \\
P(Y > t) &= 2e^{-2\lambda t} - e^{-3\lambda t}
\end{aligned}
$$

so the *cdf* of Y is given by

$$
\begin{aligned}
P(Y \leq t) &= 1 - P(Y > t) \\
&= 1 - 2e^{-2\lambda t} + e^{-3\lambda t}
\end{aligned}
$$

That's good enough.