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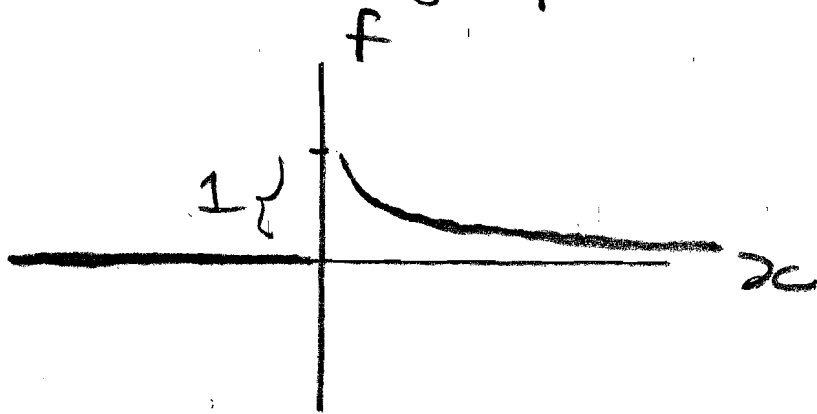
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases} \qquad (*)$$

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 $\beta = \frac{1}{\lambda}$. We will see that X is 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 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
$\qquad = $ the waiting time for
$\qquad\qquad$ 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 $0 \leq t < \infty$.

The Counting Random Variable

Now fix a time period t. So we have a "counting random variable X_t".

$$X_t = \text{\# of animals cought in the trap in time } t$$

We will choose the model
$$X_t \sim P(\lambda t) = \text{Poisson with parameter } \lambda t$$

$\angle$ We are using λ insted of α in the Poisson process.

N.B. $P(X_t = 0) = e^{-\lambda t}$ (#)

Remark

The analogue from before was

$$X_n = \# \text{ of girls in the first } n \text{ children}$$

(so we have a discrete "time period", the binomial random variable was the <u>counting random variable</u>)

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 λ

Proof

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

$\left(so \quad F(t) = P(Y \leq t) = 1 - e^{-\lambda t} \right.$

$\left. and \quad f(t) = F'(t) = \lambda e^{-\lambda t} \right)$

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

$\Longleftrightarrow$ <u>there are no animals in the trap at time t.</u>

<u>In</u> symbols this says

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

$\llcorner$ equality of events $\lrcorner$

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 r-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

$$F_r(t) = \begin{cases} 1-\left(1+\lambda+\dfrac{(\lambda t)^2}{2!}+\cdots+\dfrac{(\lambda t)^{r-1}}{(r-1)!}\right)e^{-\lambda t}, & t > 0 \\ 0, & \text{otherwise} \end{cases}$$

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

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

__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 $> t$ $\iff$ at time t there are $\leq r-1$ animals in the trap.

Since $X_t \sim P(\lambda t)$ we have

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

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

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

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

$$\text{But} \quad f_r(t) = \frac{dF_r}{dt}(t)$$

so we have to differentiate the
expression on the right-hand side

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

$$f_r(t) = -\frac{d}{dt}(e^{-\lambda t})\left(1 + \lambda t + \frac{(\lambda t)^2}{2!} + \cdots + \frac{(\lambda t)^{r-2}}{(r-2)!} + \frac{(\lambda t)^r}{(r-1)!}\right)$$

$$- e^{-\lambda t}\frac{d}{dt}\left(1 + \lambda t + \frac{(\lambda t)^2}{2!} + \cdots + \frac{(\lambda t)^{r-1}}{(r-1)!}\right)$$

$$= \lambda e^{-\lambda t}\left(1 + \lambda t + \frac{(\lambda t)^2}{2!} + \cdots + \frac{(\lambda t)^{r-2}}{(r-2)!} + \frac{(\lambda t)^{r-1}}{(r-1)!}\right)$$

$$- e^{-\lambda t}\left(\lambda + \lambda^2 t + \frac{\lambda^3 t^2}{2!} + \cdots + \frac{\lambda^{r-1}t^{r-2}}{(r-2)!}\right)$$
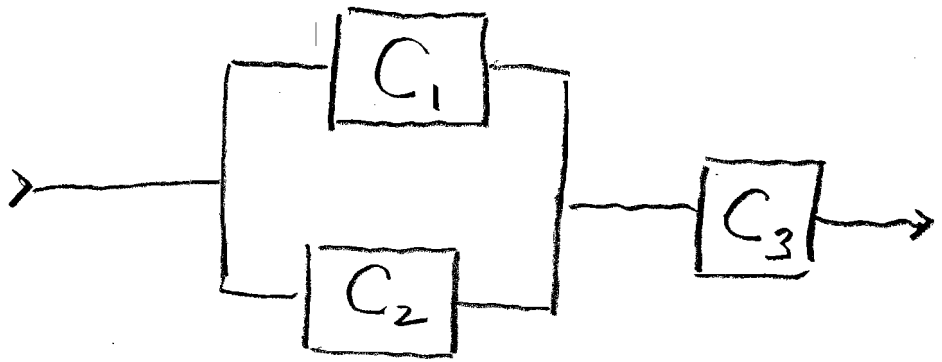
$$= e^{-\lambda t}\left(\lambda + \lambda^2 t + \frac{\lambda^3 t^2}{2!} + \cdots + \frac{\lambda^{r-1}t^{r-2}}{(r-2)!} + \frac{\lambda^r t^{r-1}}{(r-1)!}\right)$$

$$- e^{-\lambda t}\left(\lambda + \lambda^2 t + \frac{\lambda^3 t^2}{2!} + \cdots + \frac{\lambda^{r-1}t^{r-2}}{(r-1)!}\right)$$

$$= \frac{\lambda^r t^{r-1}}{(r-1)!}e^{-\lambda t}$$

□

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

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

Then X_i is exponentially distributed with parameter λ so

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

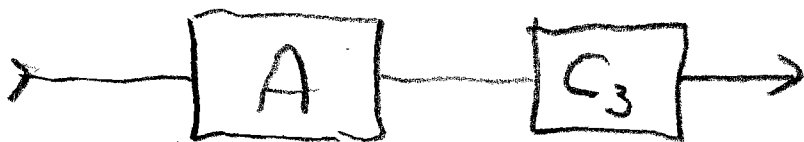$$P(X_i > t) = e^{-\lambda} \quad , \quad i = 1, 2, 3.$$

Define Y by

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

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

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

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

(the system is working at time t

$\Longrightarrow$ both A and C_3 are working at time t)

Try to get $\cap$ instead of $\cup$

— that's why I chose $(Y > t)$ on

the left.

Hence

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

by independence

$$= P(W > t) \cdot P(X_3 > t)$$

(#)

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

independent?

<u>Answer</u> Suppose $C_1, C_2, \ldots C_n$

are independent components Suppose

$A = $ a subcollection of the C_i's.
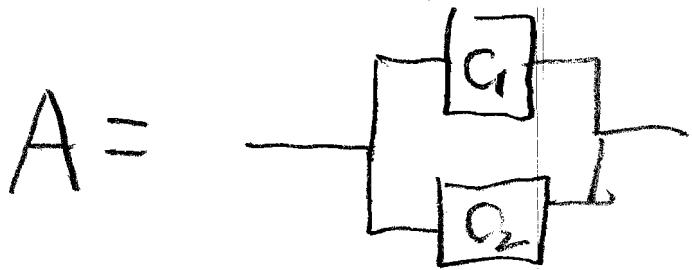
$B = $ another subcollection of the C_i's.

Then A and B are independent
$\iff$ they have no common component.

So now we need

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

$$A = $$

I should switch to
$P(W \leq t)$ to get intersections
but I want to show you why
unions give extra terms.

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

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

so

$$P(W > t) = P(X_1 > t) + P(X_2 > t)$$

extra term

$$- \boxed{P((X_1 > t) \cap (X_2 > t))}$$

$$= P(X_1 > t) + P(X_2 > t)$$

$$- P(X_1 > t) P(X_2 > t)$$

$$= e^{-\lambda t} + e^{-\lambda t} - (e^{-\lambda t})(e^{-\lambda t})$$

$$= 2e^{-\lambda t} - e^{-2\lambda t}$$

Now from (#)

$$P(Y > t) = P(W > t) P(X_3 > t)$$

$$= (2e^{-\lambda t} - e^{-2\lambda t}) e^{-\lambda t}$$

$$P(Y > t) = 2e^{-2\lambda t} - e^{-3\lambda t}$$

so the cdf of Y is given by

$$P(Y \leq t) = 1 - P(Y > t)$$

$$= 1 - 2e^{-2\lambda t} + e^{-3\lambda t}$$

That's good enough.