Lecture 27: some continuous random variables have the property that \( P(X > 0) = 1 \), so that their probability density \( f(x) \) is zero for \( x \leq 0 \). Typically:

\[
\begin{align*}
&\text{Typically:} \\
&y = f(x) \\
&0 \leq t < x
\end{align*}
\]

These are often used to model the "random time" until something happens; e.g.:
- the lifetime of an individual;
- the waiting time for the next bus;
- the time that a computer or person will take to complete a task;
- the waiting time for the emission of the next particle by a radioactive element;
- the waiting time for the next customer;
- arrival time for next internet packet;
- waiting time at the E.R.;
- duration of an electric or electronic component before it needs to be replaced;
- the length of time a patient survives after a surgery.

These are all quantities that can be described by random variables of the type (*)\footnote{math370_148_10272014}. In fact, for a r.v. \( T \) with density \( f(t) \) that is 0 for \( t < 0 \), the quantity

\[
G(t) = P(T > t) = \int_t^\infty f(x) \, dx
\]

is called "\textbf{SURVIVAL FUNCTION}".
The Exponential probability density is often used to model some (but not all) of the above random times, because it has some special property that we shall explore.

Def: for a fixed parameter $\lambda > 0$, we say that the r.v. $T$ has Exponential($\lambda$) probability density $f(t)$ if:

$$f(t) = \begin{cases} 
\lambda e^{-\lambda t} & \text{if } t \geq 0 \\
0 & \text{if } t < 0 
\end{cases}$$

Remarks:

• obviously, $\int_{-\infty}^{\infty} f(t) dt = \lambda \int_{0}^{\infty} e^{-\lambda t} dt = \frac{\lambda}{-\lambda} \left[ e^{-\lambda t} \right]_{t=0}^{t=\infty} = 1$.

• $P(a \leq T \leq b) = \lambda \int_{a}^{b} e^{-\lambda t} dt = \lambda \left[ e^{-\lambda t} \right]_{a}^{b} = e^{-\lambda a} - e^{-\lambda b}$

• Survival function: for $t > 0$

$$G(t) = P(T > t) = \lambda \int_{t}^{\infty} e^{-\lambda u} du = \lambda \left[ e^{-\lambda u} \right]_{t}^{\infty} = e^{-\lambda t}$$

• $E(T) = \lambda \int_{0}^{\infty} te^{-\lambda t} dt = \lambda \left[ -\frac{t}{\lambda} e^{-\lambda t} \right]_{0}^{\infty} - \frac{1}{\lambda} \int_{0}^{\infty} e^{-\lambda t} dt$

<table>
<thead>
<tr>
<th>$f$</th>
<th>$g$</th>
<th>$f' = 1$</th>
<th>$g = -\frac{1}{\lambda} e^{-\lambda t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{f}{\lambda} e^{-\lambda t}$</td>
<td>$\frac{-1}{\lambda} e^{-\lambda t}$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$$= \lambda \left[ e^{-\lambda t} \right]_{t=0}^{t=\infty} = \frac{1}{\lambda}$$

So: \[ E(T) = \frac{1}{\lambda} \] (this computation will be part of the next hw.)

Can prove that: \[ \text{Var}(T) = \frac{1}{\lambda^2}, \quad \text{SD}(T) = \frac{1}{\lambda} \]

(\text{this computation will be part of the next hw.})
The densities look like this:

\[
\begin{align*}
\lambda = 2 & \quad \lambda = 1 \\
\end{align*}
\]

\[SD(T) = \frac{1}{\lambda},\text{ so for}\]
A small the prob. density is more "spread out".

"Memoryless" property of the Exponential random variable
(very important!)

Theorem: if \( T \sim \text{Exponential}(\lambda) \) then, for any \( t > 0 \) \& \( s > 0 \)

\[
P(T > t + s \mid T > t) = P(T > s).
\]

Proof: \[
P(T > t + s \mid T > t) = \frac{P(T > t + s, T > t)}{P(T > t)} \quad (*)
\]

Remark: if \( A = \{T > t + s\} \), \( B = \{T > t\} \)
then \( A \subseteq B \),
(because if \( T > t + s \) then \( T + t + s > t \), so \( T > t \))
therefore \( A \cap B = A \). We get:

\[
(*) = \frac{P(T > t + s)}{P(T > t)} = \frac{G(t + s)}{G(t)} = \frac{e^{-\lambda(t + s)}}{e^{-\lambda t}} = e^{-\lambda s} = G(s) = P(T > s).
\]

Interpretation: if \( i.e. \) under the condition that we have waited for \( t \) seconds (or unit of time) for an event to happen, the probability of having to wait another \( s \) seconds is equals to the probability of waiting \( s \) seconds in the first place (as if the first \( t \) seconds had not passed).
• So we see that the exponential r.v. is especially suitable for modeling the life span of objects (such as electric or electronic components) that experience no aging effect, and whose "death" is, in a sense, a random event that has nothing to do with their age (this, of course, does not apply to living beings such as humans).

• If we're modeling the lifespan of an object as an exponential distribution, whatever the age of the thing is, the distribution of the remaining lifetime is the same as of the original lifetime distribution. (If the "average lifespan" of an electric/electronic component is 5 years and 5 years have already passed, then the remaining lifespan is, on average, still another 5 years — this obviously does not apply to humans).

• One can prove that the exponential r.v. is the only continuous r.v. with this memoryless property. You saw in a previous homework that the Geometric distribution is a discrete distribution with the memoryless property (and it is, in fact, the only one).

• \( \lambda \) is called the rate of the Exponential (\( \lambda \)) density. What is its meaning? For any \( t > 0, \Delta t > 0 \):

\[
P(T \leq t + \Delta t | T > t) = 1 - P(T > t + \Delta t | T > t) = 1 - P(T > \Delta t) = 1 - e^{-\lambda \Delta t} = 1 - [1 - \lambda \Delta t] = \lambda \Delta t, \quad \text{(for } \Delta t \text{ small)}
\]
So:

\( (*) \quad P(T \leq t + \Delta t \mid T > t) \approx \lambda \Delta t \quad \text{for } \Delta t \text{ small,} \)

i.e. the probability that the event that we are waiting for (e.g., the "death" of an electronic component) happens happens in the next \( \Delta t \) (for \( \Delta t \) small) is proportional to \( \Delta t \), and the rate \( \lambda \) is the coefficient of proportionality (the larger \( \lambda \) is, the sooner the event is likely to happen).

**Remark:** the condition \( (*) \) is actually sufficient for \( T \) to be exponentially distributed. In fact, fix \( a > 0 \) and break the interval \([0, a]\) in \( n \) small intervals, each of size \( \Delta t \):

\[
\begin{array}{cccccccccc}
 t \quad & t_0 & t_1 & t_2 & \cdots & a & t \\
\Delta t & & & & & & & & & \end{array}
\]

Assume that \( (*) \) holds, i.e. that the probability that \( T \) happens in \([i \Delta t, (i+1) \Delta t]\) is \( \approx \lambda \Delta t \).

So,

\[
P(T > a) = (1 - \lambda \Delta t)^n = (1 + \frac{-a \lambda}{n})^n \xrightarrow{n \to \infty} e^{-a} \]

which is the survival function for an exponential \((\lambda)\) probability density.

(Remember the limit: \( \lim_{n \to \infty} \left( 1 + \frac{b}{n} \right)^n = e^b \)).
**Poisson Arrival Processes** (different, but related to, Poisson Random Variables).

In a branch of mathematics called queuing theory, we have "customers" and a "server":

Customers \[\xrightarrow{ccc} S \xrightarrow{ccc} \text{Served customers}\]

Server

The Server could be the ER, and the customers are the patients that arrive at the ER.

We often model the time between two consecutive arrivals as independent and identically distributed (i.i.d.) Exponential (\(\lambda\)) random variables, where we called \(\lambda\) the "rate of arrivals", or "rate of the Poisson process".

We associate to it a "counting process" \(N(t), t \in \mathbb{R}\),

\[N(0,t] = \# \text{ of arrivals up to time } t\]

**Assumption**: the \(W_i\)'s are i.i.d. Exponential (\(\lambda\)) random variables.
The "Poisson process" may be viewed as a random distribution of the arrival points \( \{T_1, T_2, T_3, \ldots \} \) and \( N(0,t]: \mathbb{R} \to \{0,1,2,3,\ldots \} \) is a function that describes such random (ordered) set of points.

**Remark:** \( N(0,t] \) is what we call a "stochastic process", i.e. a random variable parametrized by time. In other words, for any fixed \( t > 0 \), \( N(0,t] \) is a random variable. In fact, it is a DISCRETE random variable.

**Question:** what is the distribution of \( N(0,t] \)?

(We may give the proof of the following theorem later).

**Theorem:** under the assumption that the \( W_i \)'s are iid \( \sim \text{Exponential}(\lambda) \), we have that

\[
N(0,t] \sim \text{Poisson}(\lambda t).
\]

**Remarks:**
- \( E(N(0,t]) = \lambda t \), so the number of arrivals is proportional to the rate of arrivals \( \lambda \) (for \( t \) fixed) and to the interval size \( t \) (for \( \lambda \) fixed).
- The range of \( N(0,t] \) is \( \{0,1,2,\ldots\} \), unbounded! I.e. with a (very small) probability \( > 0 \) we may have a very large number of arrivals in the interval \( (0,t] \)

More in general, for fixed \( 0 \leq s < t < \infty \), we define:

\[
N(s,t] = \# \text{ of arrivals between } s \text{ and } t,
\]

and we have \( N(s,t] \sim \text{Poisson}(\lambda(t-s)) \)

(it's a consequence of the memoryless property).