Some Continuous Distributions

The Exponential Distribution

If X has pdf $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$
($= 0$ for $x < 0$)
for some value $\lambda > 0$, we say X is an exponential rv (or we say it has an exponential distribution) and we write $X \sim Exp(\lambda)$.

We found the cdf of X is

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 . \end{cases}$$

Mean and Variance of $Exp(\lambda)$

$$EX = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^{\infty} x \lambda e^{-\lambda x}\,dx = \frac{1}{\lambda}$$

(integrate by parts once)

$$Var(X) = E(\underset{\substack{\uparrow \\ EX}}{X-\mu})^2 = \underbrace{EX^2}_{\frac{2}{\lambda^2}} - \underbrace{(EX)^2}_{\left(\frac{1}{\lambda}\right)^2} = \frac{1}{\lambda^2}$$

since

$$EX^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^{\infty} x^2 \lambda e^{-\lambda x}\,dx = \frac{2}{\lambda^2} .$$

(integrate by parts twice)

In general

$$E X^n = \int_{-\infty}^{\infty} x^n f(x) dx = \int_0^{\infty} x^n \lambda e^{-\lambda x} dx$$

integrate by parts once taking

$$u = x^n \qquad du = n x^{n-1}$$

$$dv = \lambda e^{-\lambda x} \qquad v = -e^{-\lambda x}$$

to obtain

$$= x^n(-e^{-\lambda x})\Big|_0^{\infty} - \int_0^{\infty} (-e^{-\lambda x}) n x^{n-1} dx$$

$$= 0 + n \int_0^{\infty} x^{n-1} e^{-\lambda x} dx$$

$$= \frac{n}{\lambda} \underbrace{\int_0^{\infty} x^{n-1} \lambda e^{-\lambda x} dx}_{\int_{-\infty}^{\infty} x^{n-1} f(x) dx} = \frac{n}{\lambda} E\left[X^{n-1}\right]$$

So we have $E X^n = \frac{n}{\lambda} E X^{n-1}$ for all integers $n > 0$.

Setting $n=1$ gives $E X^1 = \frac{1}{\lambda} E X^0 = \frac{1}{\lambda} \underbrace{E 1}_{=1} = \frac{1}{\lambda}$.

Setting $n=2$:

$$E X^2 = \frac{2}{\lambda} E X^1 = \frac{2}{\lambda} \cdot \frac{1}{\lambda} = \frac{2}{\lambda^2}.$$

Setting $n=3$:

$$EX^3 = \frac{3}{\lambda} EX^2 = \frac{3}{\lambda} \cdot \frac{2}{\lambda^2} = \frac{6}{\lambda^3} .$$

Setting $n=4$:

$$EX^4 = \frac{4}{\lambda} EX^3 = \frac{4}{\lambda} \cdot \frac{6}{\lambda^3} = \frac{24}{\lambda^4} .$$

And in general:

$$EX^n = \frac{n!}{\lambda^n} .$$

The exponential distribution is frequently used to model the waiting time to certain kinds of events. The time from now until...

... the next click of a geiger counter.

... your next car accident.

... the next time you win the lottery

... the next asteroid strikes the earth.

etc.

The exponential distribution is used for these situations because it has the <u>memoryless</u> property:

$$P(X > s+t \mid X > t) = P(X > s) \quad \text{for all } s,t \geq 0.$$

Think of X as the lifetime of some device which is turned on at time zero and run continuously until it fails. The memoryless property says:

If the device is still working at time t (that is, $X>t$), the probability it is still working s time units later (that is, $X>s+t$) is the same as if the device were new (that is, it was just turned on).

This statement is true for any value of t. So the memoryless property says that the probability the device will fail in the next s time units is the same no matter how long it has already been running. (The device has no memory for how long it has been running.)

<u>Proof of memoryless property</u>:

Note that $P(X>u) = \int_u^\infty \lambda e^{-\lambda x}\,dx = \left(-e^{-\lambda x}\right)\Big|_u^\infty$

$$= e^{-\lambda u} \text{ for } u \geq 0.$$

$$P(X>s+t \mid X>t) = \frac{P(\overbrace{X>s+t, X>t}^{\text{The intersection of the two events.}})}{P(X>t)}$$

Note that $X>s+t$ implies $X>t$.
Therefore $\{X>s+t\} \subset \{X>t\}$
so that $\{X>s+t, X>t\} = \{X>s+t\}$.

(In general, if A implies B
(the occurrence of A always implies the occurrence of B),
then the events satisfy $A \subset B$ so that $A \cap B = A$.)

Thus $$\frac{P(X>s+t, X>t)}{P(X>t)} = \frac{P(X>s+t)}{P(X>t)}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = \frac{e^{-\lambda s}e^{-\lambda t}}{e^{-\lambda t}} = e^{-\lambda s} = P(X>s).$$

QED

Note: Since we showed that
$$P(X>s+t \mid X>t) = \frac{P(X>s+t)}{P(X>t)} \text{ always,}$$
the memoryless property:
$$P(X>s+t \mid X>t) = P(X>s)$$
is equivalent to:
$$\frac{P(X>s+t)}{P(X>t)} = P(X>s)$$
is equivalent to
$$P(X>s+t) = P(X>s)P(X>t)$$
for $s, t \geq 0$.

Note: The exponential distribution is the only distribution which satisfies the memoryless property:

If $P(X > s+t \mid X > t) = P(X > s)$ for $s, t \geq 0$

then $X \sim \text{Exp}(\lambda)$ for some $\lambda > 0$.

The Uniform Distribution

If X has density $f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{for } \alpha < x < \beta \\ 0 & \text{otherwise,} \end{cases}$

we say that X has a uniform distribution on the interval (α, β) and write $X \sim \text{Uniform}(\alpha, \beta)$.

The function f is clearly nonnegative and integrates to 1 because the area under f is the area of the rectangle with base $= \beta - \alpha$ and height $= \frac{1}{\beta - \alpha}$.

Plot of f

The simplest uniform distribution (distn.) is Uniform$(0,1)$ which has density

$$f(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

The CDF of Uniform(α, β)

is $F(t) = \int_{-\infty}^{t} f(x)\,dx$.

If $t \leq \alpha$, then $\int_{-\infty}^{t} f(x)\,dx = \int_{-\infty}^{t} 0\,dx = 0$.

If $\alpha < t < \beta$, then $\int_{-\infty}^{t} f(x)\,dx = \int_{-\infty}^{\alpha} 0\,dx + \int_{\alpha}^{t} \frac{1}{\beta-\alpha}\,dx$

$$= 0 + \frac{1}{\beta-\alpha}\int_{\alpha}^{t} dx = \frac{t-\alpha}{\beta-\alpha}.$$

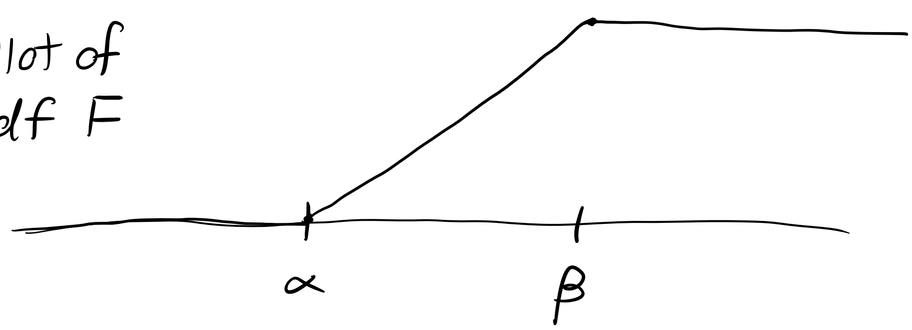
If $t \geq \beta$,

then $\int_{-\infty}^{t} f(x)\,dx = \int_{-\infty}^{\alpha} 0\,dx + \int_{\alpha}^{\beta} \frac{1}{\beta-\alpha}\,dx + \int_{\beta}^{t} 0\,dx$

$$= 0 + \int_{\alpha}^{\beta} \frac{1}{\beta-\alpha}\,dx + 0 = 1.$$

Thus $F(t) = \begin{cases} 0 & \text{for } t \leq \alpha \\ \frac{t-\alpha}{\beta-\alpha} & \text{for } \alpha < t < \beta \\ 1 & \text{for } t \geq \beta \end{cases}$

Plot of cdf F



Mean and Variance of Uniform(α, β)

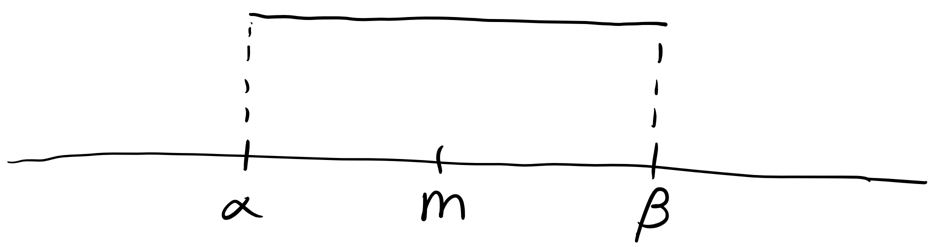
$$EX = \int_{-\infty}^{\infty} x f(x)dx = \int_{\alpha}^{\beta} x \frac{1}{\beta-\alpha} dx$$

$$= \frac{1}{\beta-\alpha} \int_{\alpha}^{\beta} x\, dx = \frac{1}{\beta-\alpha} \cdot \frac{x^2}{2}\Big|_{\alpha}^{\beta}$$

$$= \frac{1}{2} \frac{\beta^2-\alpha^2}{\beta-\alpha} = \frac{1}{2} \frac{(\beta+\alpha)(\beta-\alpha)}{\beta-\alpha} = \frac{\alpha+\beta}{2}$$

which is intuitive since $\frac{\alpha+\beta}{2}$ is the center of the interval (α, β).

The density $f(x)$ is symmetric about $m = \frac{\alpha+\beta}{2}$ meaning $f(m+x) = f(m-x)$ for all x.



When a density is symmetric about a point m, then the mean must be m because

$$EX = \int_{-\infty}^{\infty} x f(x)dx = \int_{-\infty}^{\infty} (x-m+m) f(x)dx$$

$$= \int_{-\infty}^{\infty} (x-m) f(x)dx + m \underbrace{\int_{-\infty}^{\infty} f(x)dx}_{=1}$$

Let $u = x - m$ so $x = m + u$ and $du = dx$.

$$EX = \int_{-\infty}^{\infty} u f(m+u)\, du + m$$

Thus $$EX - m = \int_{-\infty}^{\infty} u f(m+u)\, du \qquad (*)$$

$$= \int_{-\infty}^{\infty} u f(m-u)\, du \quad \left(\begin{array}{l}\text{by symmetry of} \\ f \text{ about } m\end{array}\right)$$

Now let $s = -u$, $ds = -du$.

$$= \int_{\infty}^{-\infty} (-s) f(m+s) (-ds)$$

$$= -\int_{-\infty}^{\infty} s f(m+s)\, ds \qquad (**)$$

$= 0$ by comparing $(*)$ and $(**)$ and noting that $A = -A$ implies $A = 0$.

Therefore $EX = m$.

Return to finding $Var(X)$ when $X \sim Uniform(\alpha, \beta)$.

$Var(X) = E(X - \mu)^2$ where $\mu = EX = \frac{\alpha + \beta}{2}$

$$= \int_{\alpha}^{\beta} (x - \mu)^2 \frac{1}{\beta - \alpha}\, dx$$

Let $y = x - \mu$, $dy = dx$.

$$= \int_{\alpha-\mu}^{\beta-\mu} y^2 \frac{1}{\beta-\alpha}\, dy$$

$$= \frac{1}{\beta-\alpha} \int_{-(\beta-\alpha)/2}^{(\beta-\alpha)/2} y^2\, dy$$

$$= \frac{1}{L} \int_{-L/2}^{L/2} y^2\, dy \quad \text{where } L = \beta - \alpha$$

$$= \frac{1}{L} \cdot \frac{y^3}{3} \Big|_{-L/2}^{L/2}$$

$$= \frac{1}{L} \cdot \frac{2}{3} \left(\frac{L}{2}\right)^3 = \frac{2}{3} \cdot \frac{1}{8} \cdot L^2$$

$$= \frac{1}{12} (\beta - \alpha)^2 .$$

Example: (Situation involving uniform distn.)
A bus follows a circular route. It takes an hour to complete its route and visits every place along its route at hourly intervals. Someone who does not know the schedule arrives at a bus stop at a "random" time and waits for the next bus. How long do they wait?

The waiting time X is Uniform$(0,1)$.

person arrives at a random time

Tick marks denote times bus visits that stop.

What is the probability the person waits more than 40 minutes = 2/3 hour?

$$P(X > \tfrac{2}{3}) = \int_{2/3}^{1} 1\,dx = x\Big|_{2/3}^{1} = 1 - \tfrac{2}{3}$$
$$= \tfrac{1}{3}.$$

What is the probability the person waits between 10 and 20 minutes?

$$P(\tfrac{1}{6} < X < \tfrac{2}{6}) = \int_{1/6}^{2/6} 1\,dx = \tfrac{2}{6} - \tfrac{1}{6}$$
$$= \tfrac{1}{6}.$$

<u>Example</u>: The most common use for the uniform distn. in statistics is as the distn. for P-values. If X is a random variable with cdf F, then $F(X) \sim$ Uniform$(0,1)$.

For example, suppose $X \sim Exp(\lambda)$ with cdf $F(t) = \begin{cases} 1-e^{-\lambda t} & \text{for } t \geq 0 \\ 0 & \text{for } t<0 . \end{cases}$

Define $Y = F(X) = 1-e^{-\lambda X}$.

The cdf for Y is $F_Y(t) = P(Y \leq t)$

$= P(1-e^{-\lambda X} \leq t)$

$1-e^{-\lambda X} \leq t$ iff $1-t \leq e^{-\lambda X}$

iff $\log(1-t) \leq -\lambda X$

iff $X \leq -\frac{1}{\lambda}\log(1-t)$.

Thus
$$
\begin{aligned}
F_Y(t) &= P(1-e^{-\lambda X} \leq t) \\
&= P(X \leq -\tfrac{1}{\lambda}\log(1-t)) \\
&= \underset{\substack{\uparrow \\ \text{cdf of } X}}{F}(-\tfrac{1}{\lambda}\log(1-t)) \\
&= 1-e^{-\lambda[-\frac{1}{\lambda}\log(1-t)]} \\
&= 1-e^{\log(1-t)} \\
&= 1-(1-t) = t .
\end{aligned}
$$

$F_Y(t) = t$ for $0<t<1$ is the cdf for the Uniform(0,1) distn. Therefore $Y \sim$ Uniform(0,1).