Continuous Distributions, Exponential Distribution

Sta 111

Colin Rundel

May 23, 2014

Cumulative Distribution Function

We have already seen a variety of problems where we find $P(X \leq x)$ or $P(X > x)$ etc. The former is given a special name - the cumulative distribution function.

If X is discrete with probability mass function $f(x)$ then

$$P(X \leq x) = F(x) = \sum_{z=-\infty}^{x} f(z)$$

If X is continuous with probability density function $f(x)$ then

$$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(z)dz$$

CDF is defined for for all $-\infty < x < \infty$ and follows the following rules:

- $\lim_{x\to-\infty} F(x) = 0$
- $\lim_{x\to\infty} F(x) = 1$
- $x < y \Rightarrow F(x) < F(y)$

Binomial CDF

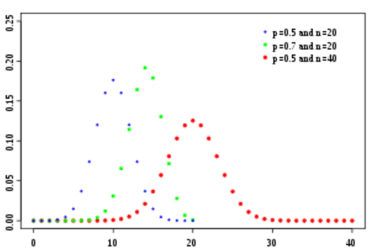
Let $X \sim \text{Binom}(n, p)$ then

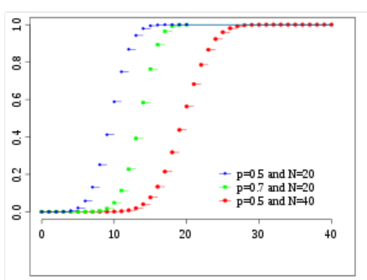
Probability Mass Function

$$P(X = k) = f(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Cumulative Density Function

$$P(X \leq x) = F(x) = \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k (1-p)^{n-k}$$





Probability Density Function

For a continuous probability distribution

$$P(X = x) = 0 \text{ for all } x$$

As such we define the probability density function to be

$$f_X(x) = \lim_{\epsilon\to 0} P(X \in [x, x+\epsilon])/\epsilon$$

A pdf is defined for for all $-\infty < x < \infty$ and follows the following rules:

- $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
- $\int_{-\infty}^{x} f_X(t)\, dt = F_X(x) \;\Leftrightarrow\; f_X(x) = \frac{d}{dx} F_X(x)$
- $f_X(x) \geq 0$ for all x

Uniform CDF

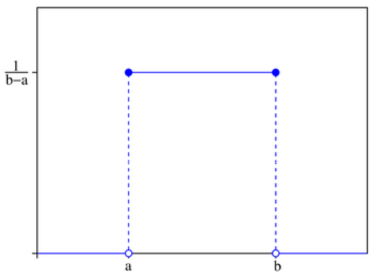
Let $X \sim \text{Unif}(a, b)$ then

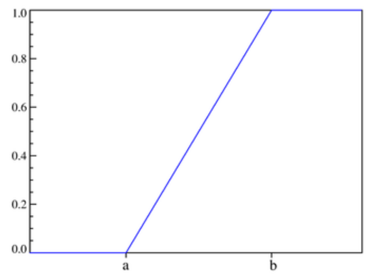
Probability Mass Function

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

Cumulative Density Function

$$F(x) = \begin{cases} 0 & \text{for } x \le a \\ \frac{x-a}{b-a} & \text{for } x \in [a, b] \\ 1 & \text{for } x \ge b \end{cases}$$





Normal CDF

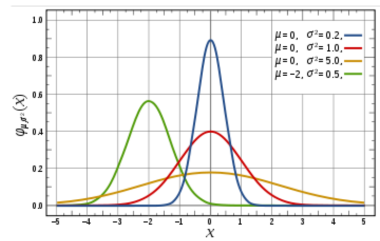
Let $X \sim \mathcal{N}(\mu, \sigma^2)$ then

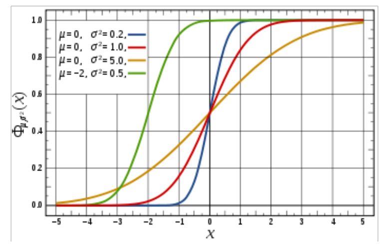
Probability Mass Function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Cumulative Density Function

$$F(x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx$$





Probability density function weirdness

Let $X = U^2$ where $U \sim \text{Unif}(0, 1)$ what are the cdf and pdf of X?

$$F(x) = P(X \le x) = P(U^2 \le x) = P(U \le \sqrt{x}) = \begin{cases} 0 & \text{if } x < 0 \\ \sqrt{x} & \text{if } 0 \le x \le 1 \\ 1 & \text{if } x > 1 \end{cases}$$

$$f(x) = \frac{d}{dx} F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2\sqrt{x}} & \text{if } 0 \le x \le 1 \\ 0 & \text{if } x > 1 \end{cases}$$

What is the value of $f(x)$ when $x = 0 + \epsilon$? What is $P(X = 0 + \epsilon)$?

Exponential Distribution

We derive the Exponential distribution by thinking of it as a RV that describes the waiting time between events which occur continuously with the rate λ.

λ here has the same meaning as in the Poisson distribution, it is the expected number of events in a given unit of time.

Let us consider one such unit of time, we expect that there will be λ events in this time span. If we subdivide that unit of time into n subinterval then the probability that one of the events falls with a certain subinterval should be approximately λ/n.

Exponential Distribution, cont.

Let $X \sim \text{Exp}(\lambda)$, we start by examining $P(X \le b)$ where b is a positive integer. This is in essence asking, what is the probability that we do not have to wait longer than b units of time before the first event occurs.

Since we have divided each unit of time up into n subdivisions, this is the same as asking what is the probability that the event occurs in the first nb sub-intervals.

Since we have the (approximate) probability of the event for each subinterval we can model this probability with a Geometric random variable Y with $p = \lambda/n$.

$$P(X \le b) \approx P(Y \le nb) = \sum_{k=0}^{bn-1} P(Y = k) = \sum_{k=0}^{bn-1} \left(1 - \frac{\lambda}{n}\right)^k \frac{\lambda}{n}$$

Exponential Distribution, cont.

From calculus remember that:

$$\sum_{k=0}^{m} a^k = \frac{1 - a^{m+1}}{1 - a}$$

Therefore,

Exponential Distribution, cont.

In this case we have the CDF but not the PDF, how do we get the PDF?

Exponential Distribution, cont.

Let X be a random variable that reflects the time between events which occur continuously with a rate λ, $X \sim \text{Exp}(\lambda)$

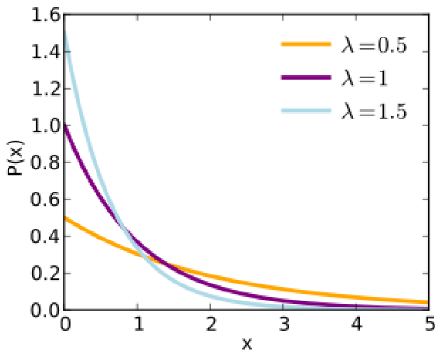
$$f(x|\lambda) = \lambda e^{-\lambda x}$$
$$P(X \le x) = F(x|\lambda) = 1 - e^{-\lambda x}$$

$$M_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-1}$$

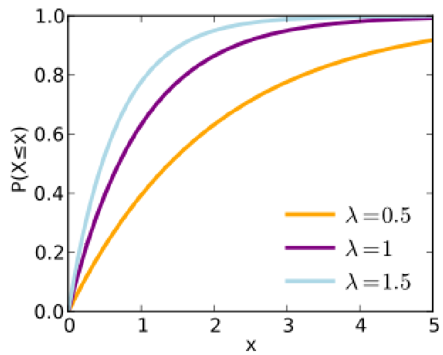
$$E(X) = \lambda^{-1}$$
$$Var(X) = \lambda^{-2}$$
$$\text{Median}(X) = \frac{\log 2}{\lambda}$$

Exponential Distribution, cont.

Probability density function:



Cumulative distribution function:



Exponential Distribution - Memoryless Property

Let $X \sim \text{Exp}(\lambda)$ (assume λ has units of events/min) then if we have waited s minutes without observing an event what is the probability that an event occurs in the next t minutes?

Exponential Distribution - Example

Strontium 90 is a radioactive component of fallout from nuclear explosions. The halflife of Strontium 90 is 28 years and the decay of an individual atom can be modeled by an exponential random variable.

(a) What is the decay rate λ?

(b) What is the average lifetime of a Strontium 90 atom?

(c) What is the probability that a Strontium 90 atom survives at least 50 years?

(d) What is the probability that a Strontium 90 survives at least 75 years given it has survived at least 25 years?