Lecture 10: Central Limit Theorem and CDFs

Sta230 / Mth 230

Colin Rundel

February 25, 2014

Moments

Raw moment:

$$\mu'_n = E(X^n)$$

Central moment:

$$\mu_n = E[(X-\mu)^2]$$

Normalized / Standardized moment:

$$\frac{\mu_n}{\sigma^n}$$

Moment Generating Function

The moment generating function of a random variable X is defined for all real values of t by

$$M_X(t) = E[e^{tX}] = \begin{cases} \sum_x e^{tx} P(X=x) & \text{If } X \text{ is discrete} \\ \int_x e^{tx} P(X=x)dx & \text{If } X \text{ is continuous} \end{cases}$$

This is called the moment generating function because we can obtain the raw moments of X by successively differentiating $M_X(t)$ and evaluating at $t=0$.

$$M_X(0) = E[e^0] = 1 = \mu'_0$$

$$M'_X(t) = \frac{d}{dt}E[e^{tX}] = E\left[\frac{d}{dt}e^{tX}\right] = E[Xe^{tX}]$$
$$M'_X(0) = E[Xe^0] = E[X] = \mu'_1$$

$$M''_X(t) = \frac{d}{dt}M'_X(t) = \frac{d}{dt}E[Xe^{tX}] = E\left[\frac{d}{dt}(Xe^{tX})\right] = E[X^2e^{tX}]$$
$$M''_X(0) = E[X^2e^0] = E[X^2] = \mu'_2$$

Moment Generating Function - Properties

If X and Y are independent random variables then the moment generating function for the distribution of $X+Y$ is

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tX}e^{tY}] = E[e^{tX}]E[e^{tY}] = M_X(t)M_Y(t)$$

Similarly, the moment generating function for S_n, the sum of iid random variables $X_1, X_2, \ldots, X_n$ is

$$M_{S_n}(t) = [M_{X_i}(t)]^n$$

Moment Generating Function - Unit Normal

Let $Z \sim \mathcal{N}(0,1)$ then

Moment Generating Function - Unit Normal, cont.

Central Limit Theorem

Let $X_1, \ldots, X_n$ be a sequence of independent and identically distributed random variables each having mean μ and variance σ^2. Then the distribution of

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

tends to the unit normal as $n \to \infty$.

That is, for $-\infty < a < \infty$,

$$P\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-x^2/2} dx = \Phi(a) \text{ as } n \to \infty$$

Sketch of Proof

Proposition

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables and $S_n = X_1 + \cdots + X_n$. The distribution of S_n is given by the distribution function f_{S_n} which has a moment generating function M_{S_n} with $n \geq 1$.

Let Z being a random variable with distribution function f_Z and moment generating function M_Z.

If $M_{S_n}(t) \to M_Z(t)$ for all t, then $f_{S_n}(t) \to f_Z(t)$ for all t at which $f_Z(t)$ is continuous.

We can prove the CLT by letting $Z \sim \mathcal{N}(0,1)$, $M_Z(t) = e^{t^2/2}$ and then showing for any S_n that $M_{S_n/\sqrt{n}} \to e^{t^2/2}$ as $n \to \infty$.

Proof of the CLT

Some simplifying assumptions and notation:

- $E(X_i) = 0$
- $Var(X_i) = 1$
- $M_{X_i}(t)$ exists and is finite
- $L(t) = \log M(t)$

Also, remember L'Hospital's Rule:

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}$$

Proof of the CLT, cont.

The moment generating function of $X_i/\sqrt{n}$ is given by

$$M_{X_i/\sqrt{n}}(t) = E\left[\exp\left(\frac{tX_i}{\sqrt{n}}\right)\right] = M_{X_i}\left(\frac{t}{\sqrt{n}}\right)$$

and this the moment generating function of $S_n/\sqrt{n} = \sum_{i=1}^{n} X_i/\sqrt{n}$ is given by

$$M_{S_n/\sqrt{n}}(t) = \left[M_{X_i}\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

Therefore in order to show $M_{S_n/\sqrt{n}} \to M_Z(t)$ we need to show

$$\left[M_{X_i}\left(\frac{t}{\sqrt{n}}\right)\right]^n \to e^{t^2/2}$$

Proof of the CLT, cont.

Proof of the CLT, cont.

Proof of the CLT, Final Comments

The preceding proof assumes that $E(X_i) = 0$ and $Var(X_i) = 1$.

We can generalize this result to any collection of random variables Y_i by considering the standardized form $Y_i^* = (Y_i - \mu)/\sigma$.

$$\begin{aligned} \frac{Y_1 + \cdots + Y_n - n\mu}{\sigma\sqrt{n}} &= \left(\frac{Y_1 - \mu}{\sigma} + \cdots + \frac{Y_n - \mu}{\sigma}\right) \Big/ \sqrt{n} \\ &= \left(Y_1^* + \cdots + Y_n^*\right) / \sqrt{n} \end{aligned}$$

$$\begin{aligned} E(Y_i^*) &= 0 \\ Var(Y_i^*) &= 1 \end{aligned}$$

Cumulative Distribution Function

We have already seen a variety of problems where we find $P(X <= x)$ 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

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

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

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

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

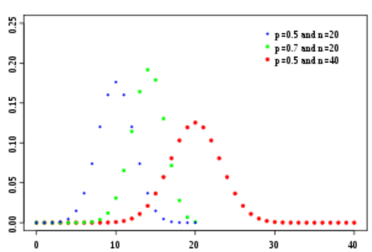
Let $X \sim \text{Binom}(n, p)$ then

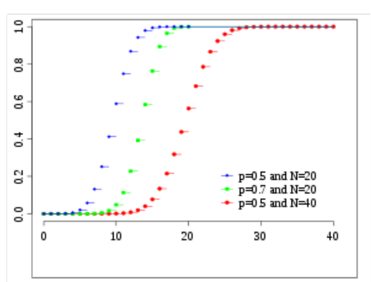
Probability Mass Function

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

Cumulative Density Function

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





Uniform CDF

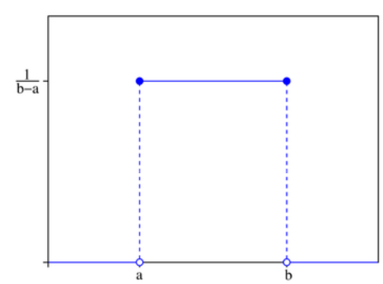
Let $X \sim \text{Unif}(a, b)$ then

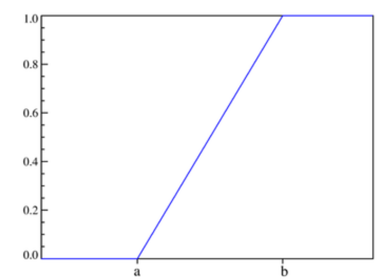
Probability Mass Function

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

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

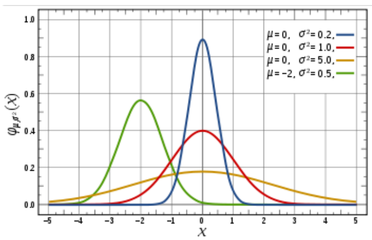
Let $X \sim \mathcal{N}(\mu, \sigma^2)$ then

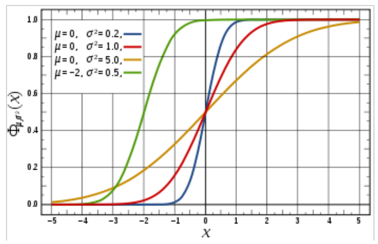
Probability Mass Function

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

Cumulative Density Function

$$F(x) = \Phi(x)$$





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 - 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?