Stat 134 lec 21

Last time

Sec 4.1 Continuous distributions

① Probability density function $\quad f(x) \geq 0$ for $x \in X$

and $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

$$P(X=a) = \int_a^a f(x)\,dx = \boxed{0}$$

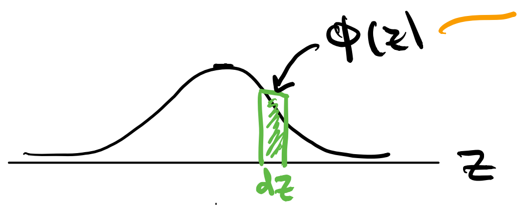
so $P(X \geq a) = P(X > a)$

ex $Z \sim N(0,1)$

$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ density function

$P(Z<1) = \Phi(1)$ CDF

units: probability / unit length



③ Change of scale of RV X — $aX+b$

ex $X \sim N(\mu, \sigma^2)$

Find $P(X>a)$

Make change of scale $Z = \frac{X-\mu}{\sigma} = \frac{1}{\sigma}X + -\frac{\mu}{\sigma}$

$\Rightarrow X = \sigma Z + \mu$

$$P(X>a) = P(\sigma Z + \mu > a) = P\left(Z > \frac{a-\mu}{\sigma}\right)$$

$$= \boxed{1 - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

Today

Sec 4.5 Cumulative Distribution function (CDF)

Sec 4.2 Exponential Distribution

Sec 4.5 The Cumulative Distribution function (CDF).

Let X be a RV. The CDF of X is

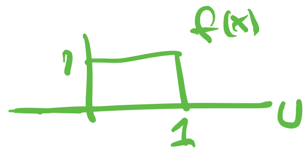
$F(x) = P(X \le x)$. — a number between 0 and 1.

If $f(x)$ is the density of X,

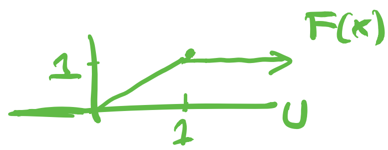
$$P(X \le x) = \int_{-\infty}^{x} f(x)\,dx$$

ex $X \sim \text{Unif}(0,1)$

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



$$F(x) = \int_{0}^{x} 1\,dx = x$$



By the FTC, $F'(x) = f(x)$.

Consequently a density function and cdf are equivalent descriptions of a RV.

Sec 4.2 Exponential and Gamma Distribution

Memoryless Property of the geometric distribution

$X \sim Geom(p)$

X = # trials until your first success.

$P(X=k) = q^{k-1}p$

$P(X>k) = q^k$

Note that $P(X=k) = P(X>k-1) - P(X>k)$

$$= q^{k-1} - q^{k} = q^{k-1}(1-q) = q^{k-1}p$$

So $P(X>k) = q^k \Rightarrow X$ is $Geom(p)$.

Theorem

The Geometric distribution is the only discrete distribution with values $1, 2, 3, \ldots$ having the memoryless property

$$P(X > k+j \mid X > j) = P(X > k)$$

In words:

If it takes you more than $j=10$ coin tosses to get your first heads, what is the chance it will take you more than $k+j=13$ coin tosses to get your first heads?

Answer: The same as the unconditional chance of needing more than $k=3$ coin tosses to get your first heads.

Pf/ Let $X \sim \text{Geom}(p)$, $X = 1, 2, 3, \ldots$

$$P(X > k+j \mid X > j) = \frac{P(X > k+j, X > j)}{P(X > j)}$$

$$= \frac{P(X > k+j)}{P(X > j)} = \frac{q^{k+j}}{q^j} = q^k$$

$$= P(X > k).$$

Conversely,

For positive integers k, j,

Suppose $P(X > k+j \mid X > j) = P(X > k)$.

$$\Vert$$

$$\frac{P(X > k+j)}{P(X > j)}$$

$\Rightarrow$ $P(X > k+j) = P(X > j) P(X > k)$

let $q = P(X > 1)$

we show that $P(X > j) = q^j$ by induction.

base case:

$P(X > 1) = q^1$ ✓

Assume $P(X > j-1) = q^{j-1}$

$P(X > j) = P(X > \underbrace{j-1}_{} + 1) = P(X > j-1) P(X > 1) = q^j$

$\Rightarrow X \sim \text{Geom}(p)$.

$\square$

Only 2 distributions are memoryless

For discrete $(X: 1, 2, 3, \ldots)$ — Geometric

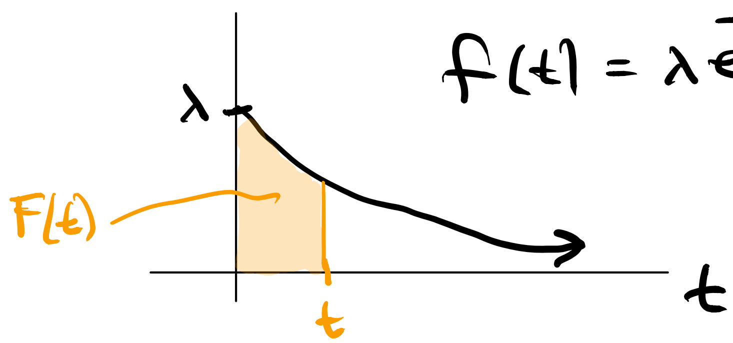
For continuous $(T > 0)$ — exponential

Proof similar to geometric

Exponential distribution

Def^n A random time T has exponential distribution with rate $\lambda > 0$

$T \sim \text{expon}(\lambda)$, if T has density $f(t) = \lambda e^{-\lambda t} \quad t \geq 0$



T = time until your first success where λ is rate of success.

Ex T = time until a lightbulb burns out.

CDF and survival function

$T \sim \text{expon}(\lambda)$

Compute the CDF $F(t) = P(T \leq t)$

$$F(t) = \int_0^t f(s)\,ds = \int_0^t \lambda e^{-\lambda s}\,ds = \frac{\lambda e^{-\lambda s}}{-\lambda}\Big|_0^t$$

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

$P(T>t)$ is called the survival function

$$\boxed{P(T>t) = e^{-\lambda t}}$$

Since $F(t) = 1 - P(T>t)$ and $f(t) = F'(t)$,

$P(T>t) = e^{-\lambda t} \Rightarrow T \sim \text{expon}(\lambda)$

Hence, $T \sim \text{expon}(\lambda)$ iff $P(T>t) = e^{-\lambda t}$.

Can we think of geometric as a discrete version of exponential? – yes.

let $X \sim \text{Geom}(p)$ $\quad X = 1, 2, 3, \ldots$

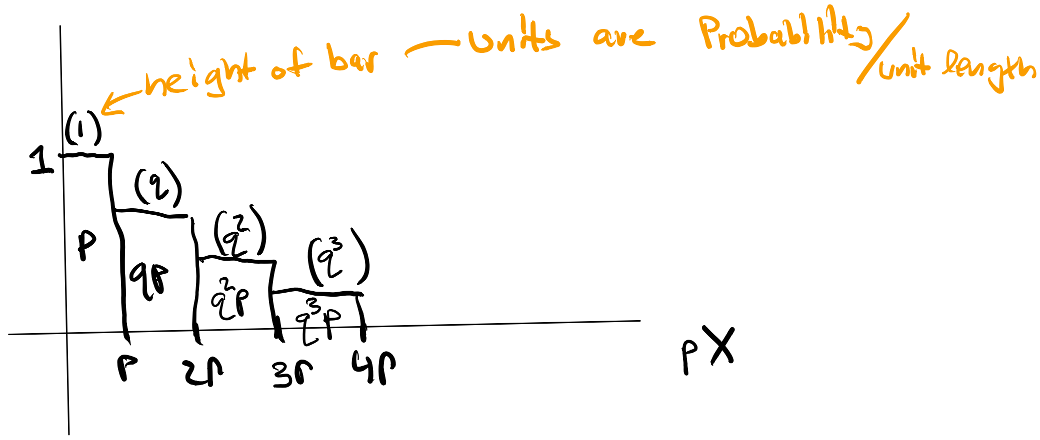
let pX be a change of scale of X

$$pX = p, 2p, 3p, \ldots$$

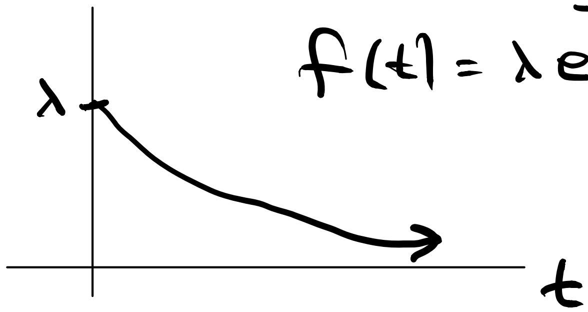
we will show that $pX \to \text{expon}(1)$ as $p \to 0$.

$$P(pX = p) = P(X=1) = p$$
$$P(pX = 2p) = P(X=2) = qp$$
$$P(pX = 3p) = P(X=3) = q^2p$$
$$\vdots$$

Draw the prob mass function for pX:



Compare this with the density of expon(λ).



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

Review Taylor series for

$$e^{x} = 1 + x + \frac{x^2}{2!} + \cdots \approx 1 + x \quad \text{for small } x.$$

$$\Rightarrow \boxed{e^{-x} \approx 1 - x} \quad \text{for small } x$$

The height of bar at time $t = kp$ sec

is $q^{k} = (1-p)^{k} \underset{p \text{ small}}{\approx} (e^{-p})^{k} = e^{-kp} = e^{-t}$

as $p \to 0$ $\quad f(t) = 1\, e^{-t}$.

$\Rightarrow pX \sim$ expon$(\lambda = 1)$ for very small p.

The height of the first bar is always 1 (no matter what p is).

The units of the height is probability/$\underbrace{\text{unit length}}_{\substack{k \\ p}}$.

As $p \to 0$, the height of the first bar stays at 1 and we get $\lambda = 1$.

λ is called the "instantaneous death rate".

It is the y intercept in the graph of the density function.

Interpretation: λ is the prob of successes during an infinitesimally small time interval.

In summary: For $X \sim \text{Geom}(p)$, $pX \to \text{expon}(1)$ as $p \to 0$.

<u>ex</u>

. Suppose X follows an exponential distribution with parameter $\lambda = 5$. Let $Y = \lceil X \rceil$. $\lceil X \rceil$ is called the ceiling function of X: it gives the smallest integer $Y \geq X$. Example: $\lceil 3.14 \rceil = 4$. Identify the distribution Y. Please show all of your work below. (Hint: Can you find $P(Y = y)$?)

<u>soln</u>

$y \in \mathbb{N}$

$$P(Y=y) = P(X \in (y-1, y))$$
$$= P(X > y-1) - P(X > y)$$
$$= e^{-5(y-1)} - e^{-5y}$$

$$= e^{-5y+5} - e^{-5y}$$

$$= e^{-5y}\left(e^{5}-1\right)$$

$$= e^{-5y}\, e^{5}\left(1-e^{-5}\right)$$

$$= \underbrace{e^{-5(y-1)}}_{q}\;\underbrace{\left(1-e^{-5}\right)}_{p}$$

$$= q^{y-1}p$$

$$\rightarrow \boxed{Y \sim \mathrm{Geom}\left(1-e^{-5}\right)}$$