317 2016-02-04

Note Title 2016-02-04

Some common discrete distributions

Bernoulli(p)

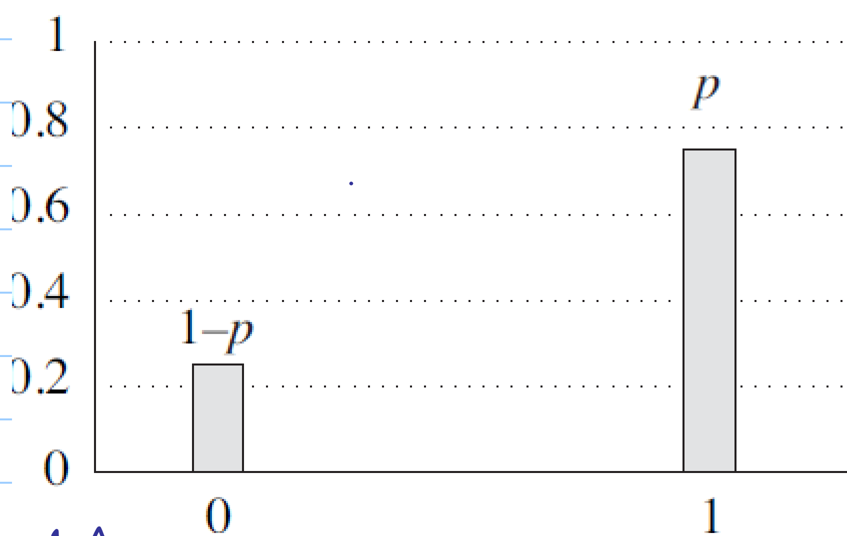
This corresponds to an experiment in which a single p-biased coin is tossed.

The probability mass function (pmf) for a Bernoulli(p) random variable X (indicated $X \sim$ Bernoulli(p)) is as follows

$$P_X(1) = p \quad P\{X=1\} \qquad X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

$$P_X(0) = 1-p \quad P\{X=0\}$$

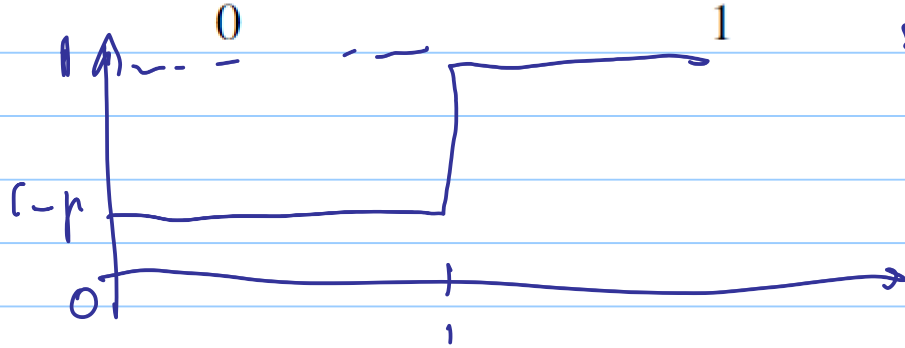
Probability mass function of the (a) Bernoulli(p) r.v.



$P\{X=0\} = 1-p$

$P\{X=1\} = p$

cdf



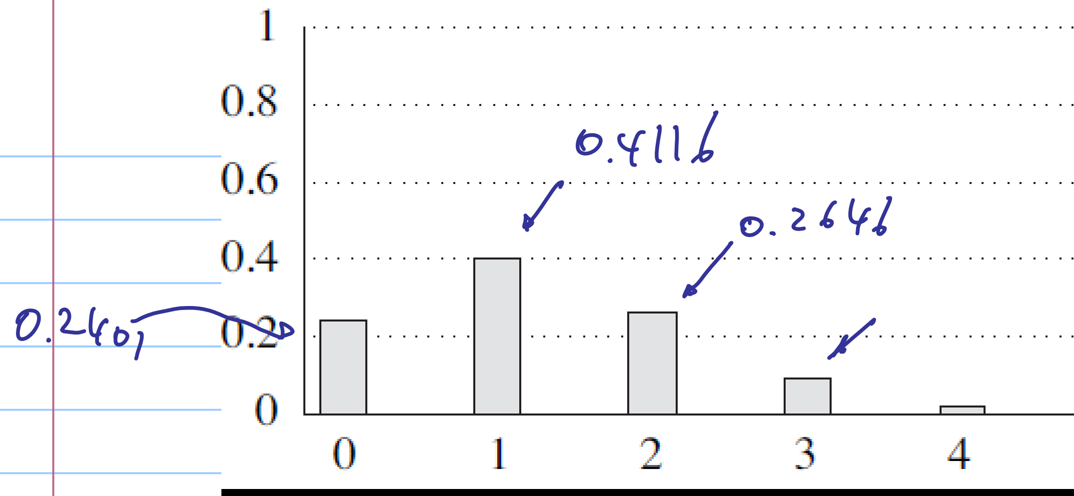
$F_X(a) = P\{X \leq a\}$

Binomial(n,p) $\quad X$ Two parameters

A binomial(n,p) r.v. X counts the number of heads on tossing a p-biased coin n times.

X may take values in the integers between 0 and n.

The p.m.f. for X is: $P_X(i) = P\{X=i\} = \binom{n}{i} p^i (1-p)^{n-i}$, for $0 \le i \le n$



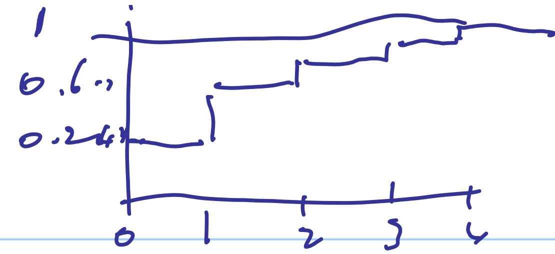
$$P\{X=2\} = \binom{4}{2} 0.3^2 \times 0.7^2 = \frac{4!}{2!\,2!} 0.3^2 \times 0.7^2 = \frac{24}{2 \cdot 2} \times 0.3^2 \times 0.7^2 = 6 \times 0.3^2 \times 0.7^2 = 0.2646$$

The probability mass function of the (a) Binomial$(4, 0.3)$ random variable X.

$$P\{X=0\} = \binom{4}{0} 0.3^0 (0.7)^4 = 0.3^0 \times 0.7^4 = 0.7^4 = 0.2401$$

$$P\{X=1\} = \binom{4}{1} 0.3^1 (0.7)^3 = 4 \times 0.3 \times 0.7^3 = 0.4116$$

c.d.f of



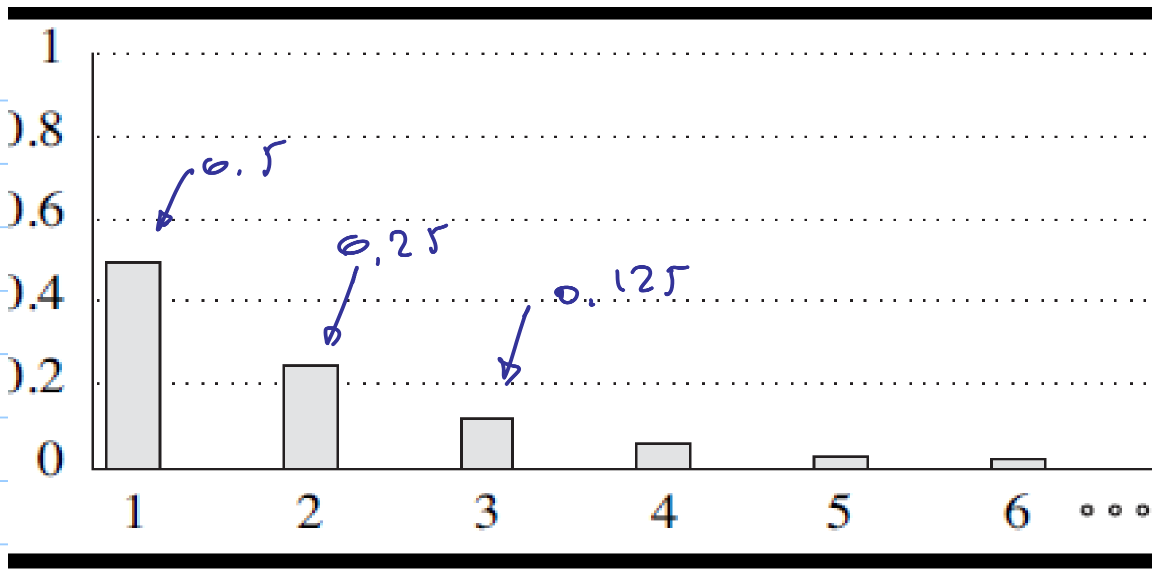
of $X \sim \text{Binomial}(4, 0.7)$

Geometric (p) distribution.

$X \sim \text{Geometric}(p)$ counts the number of tosses of a p-biased coin until the first success.

X may take any positive integer value.

$p_X(i) = P\{X = i\} = (1-p)^{i-1}\, p$, where $i = 1, 2, \cdots$



pmf of $X \sim \text{Geometric}(0.5)$

$P\{X=1\} = P_X(1) = 0.5$

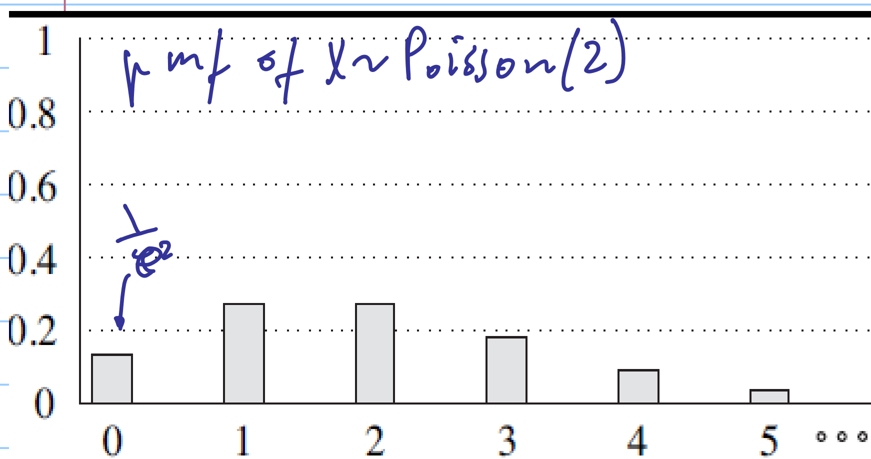
$P\{X=2\} = P_X(2) = (1-0.5)^1\, 0.5^1 = 0.5^2 = 0.25$

$P\{X=3\} = P_X(3) = (1-0.5)^2\, 0.5 = 0.25 \times 0.5 = 0.125$

The Poisson distribution. The pmf of $X \sim \text{Poisson}(\lambda)$

(It will used to model the number of arrivals to a website or router per unit time; "arrival rate is Poisson distributed" This is a misleading assumption that is often used.)

$$p_X(i) = P\{X = i\} = \frac{e^{-\lambda}\lambda^{i}}{i!}, \quad i = 0, 1, 2, \ldots$$



$$p_X(0) = \frac{e^{-2}\cdot 2^{0}}{0!} = \frac{e^{-2}\cdot 1}{1} = \frac{1}{e^{2}}$$

$$p_X(1) = \frac{e^{-2}\cdot 2^{1}}{1!} = 2e^{-2} = \frac{2}{e^{2}}$$

$$p_X(2) = \frac{e^{-2}\cdot 2^{2}}{2!} = \frac{4}{2}\frac{1}{e^{2}} = \frac{2}{e^{2}}$$

Binomial (n, p) can be approximated by Poisson(np) for large n.

For example, if you have a large enough farm of n disks, you may approximate the failure probability X, which is exactly $X \sim$ Binomial(n, p), by

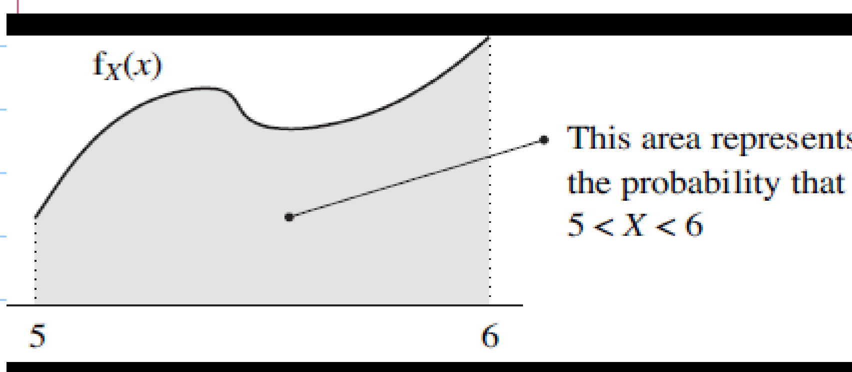
$X \sim$ Poisson(np).

Continuous r.v.s

Def: The probability density function of a continuous r.v. X is a non-negative fcn $f_X(\cdot)$, where

$$P\{a \le X \le b\} = \int_a^b f_X(x)\,dx, \text{ where } \int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

lower-case x — dummy variable (you may use y, if you prefer)



Example

$$P\{5 \le X \le 6\} = \int_5^6 f_X(x)\,dx$$

Example 1

$$f_X(x) = \begin{cases} .5x^{-.5} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Is this a pdf?

Is it non-negative? $\underline{\text{no}}$

Does it integrate to 1?

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^1 0.5\,x^{-0.5}\,dx = 0.5\int_0^1 x^{-\frac{1}{2}}\,dx = \cancel{0.5}\,\frac{x^{\frac{1}{2}}}{\cancel{\frac{1}{2}}}\Bigg|_0^1 =$$

$$= x^{\frac{1}{2}}\Big|_0^1 = 1^{\frac{1}{2}} - 0^{\frac{1}{2}} = 1 \quad \checkmark$$

Another one

$$f_X(x) = \begin{cases} 2x^{-2} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

nonnegative ✓

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^1 2x^{-2}\,dx = 2\int_0^1 x^{-2}\,dx = 2\left.\frac{x^{-1}}{-1}\right|_0^1 = \left.-\frac{2}{x}\right|_0^1 = -2-(-\infty)$$

diverges!

$f_X(x)$ is not a pdf, b/c it does not integrate to 1.

$$f_X(x) = \begin{cases} x^{-2}, & 1 < x < \infty \\ 0 & \text{otherwise} \end{cases}$$ nonnegative ✓

$$\int_{-\infty}^{\infty} x^{-2}\, dx = \frac{x^{-1}}{-1}\Bigg|_1^{\infty} = -\frac{1}{x}\Bigg|_1^{\infty} = 0 \oplus (-1) = 1 \checkmark$$

The cumulative distribution function (cdf) of an rv X is

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

Also, $\bar{F}_X(a) = 1 - F_X(a) = P\{X > a\}$

$$f_X(x) = \frac{d}{dx}\int_{-\infty}^{x} f(t)\,dt = \frac{d}{dx}F_X(x)$$