Geometric distribution.

$X \sim \text{Geom}(p)$ if

$$p_X(x) = \begin{cases} p(1-p)^{x-1} & x \in \{1, 2, \dots\} \\ 0 & \text{o/w.} \end{cases}$$

Interpret this as the number of independent $\text{Ber}(p)$ trials necessary to see the first success.

(Intuitively) $P(\text{first success happens on trial } x)$

$= P(\text{first } x-1 \text{ trials fail } \& \text{ trial } x \text{ is a success})$

$= (1-p)^{x-1} p.$

But: what is the probability space?

Want to take it to be

$\Omega = \{\text{success, failure}\}^{\mathbb{N}}$

$X(S, S, S, F, F, S, \dots) = 1.$

Problem: Ω is uncountable.

So we'll see how to work with it rigorously next month.

But for now, can check that

$\sum_x p_X(x) = 1.$ So could take $\Omega = \{1, 2, \dots\}.$

$p(k) = p_X(k).$

Negative binomial distribution

We say that X has the negative binomial distn with parameters $r \in \mathbb{N}$ and $p \in [0,1]$ if $p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$, $k \geq r$, $k \in \mathbb{N}$

which is the distribution of the first time we have seen r successes.

<u>Note:</u> $p_X(k) = p \cdot \underbrace{P\left(\begin{array}{c}\text{exactly } r-1 \text{ successes} \\ \text{in the first } k-1 \text{ trials}\end{array}\right)}_{\binom{k-1}{r-1} p^{r-1} (1-p)^{k-r}}$.

Theorem (about cdfs)

If F is the cdf of some r.v. (i.e. $F(x) = P(X \le x)$) then F has the following properties:

(a) $\lim_{x \to \infty} F(x) = 1$

(b) $\lim_{x \to -\infty} F(x) = 0$

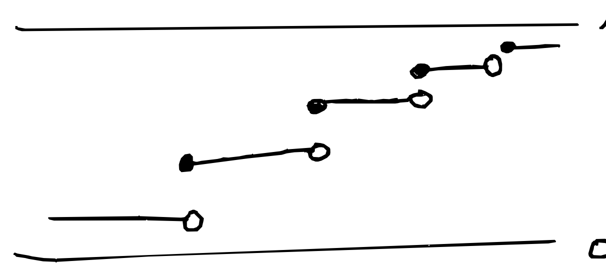
(c) F is non-decreasing

(d) F is right-continuous with left limits (RCLL) (càdlàg)

(e) $P(X > x) = 1 - F(x)$

(f) $P(x < X \le y) = F(y) - F(x)$

(g) $P(X = x) = F(x) - \lim_{y \uparrow x} F(y)$ size of the jump



Lemma Let $A_1 \subseteq A_2 \subseteq \cdots$ be a sequence of events. and $A = \bigcup_{i=1}^{\infty} A_i$. Then

$$P(A) = \lim_{i \to \infty} P(A_i).$$

Also: Let $B_1 \supseteq B_2 \supseteq \cdots$, $B = \bigcap_{i=1}^{\infty} B_i$. Then

$$P(B) = \lim_{i \to \infty} P(B_i).$$

Proof. We can write

$$A = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup \cdots$$

and the events on the r.h.s. union are disjoint.

$$\text{So } P(A) = P(A_1) + \sum_{i=1}^{\infty} P(A_{i+1} \setminus A_i)$$

$$= P(A_1) + \lim_{n\to\infty} \sum_{i=1}^{n} P(A_{i+1} \setminus A_i)$$

because $A_i \subseteq A_{i+1}$.

$$= P(A_1) + \lim_{n\to\infty} \sum_{i=1}^{n} \left(P(A_{i+1}) - P(A_i)\right)$$

$$= P(A_1) + \lim_{n\to\infty} P(A_{n+1}) - P(A_1)$$

$$= \lim_{n\to\infty} P(A_n). \qquad \square$$

For the statement with the B's write $A_i = B_i^c$

$$A = B^c$$

and use the statement for the A's.

Now we can use this to prove the theorem.

(a). Consider any seq. $x_n \uparrow \infty$

Define $A_n = \{X \le x_n\}$.

Then $\bigcup_{n \ge 1} A_n = \Omega$.

So Lemma $\Rightarrow$ $1 = P(\Omega) = \lim_{n \to \infty} P(A_n)$

$= \lim_{n \to \infty} F(x_n)$.

Since this holds for every seq $x_n \uparrow \infty$

This means that $\lim_{x \uparrow \infty} F(x) = 1$.

(b) Similar.

(c). if $x \le y$ then $\{X \le x\} \subseteq \{X \le y\}$

so $F(x) = P(X \le x) \le P(X \le y) = F(y)$.

(d) For any sequence $x_n \downarrow x$ we have

$$\{X \le x\} = \bigcap_{n=1}^{\infty} \{X \le x_n\}.$$

So Lemma $\Rightarrow$

$$P(X \le x) = \lim_{n \to \infty} P(X \le x_n)$$

so $F(x) = \lim_{n \to \infty} F(x_n)$.

Since this is true for every sequence $x_n \downarrow x$

$$F(x) = \lim_{y \downarrow x} F(y)$$

i.e. F is right-continuous

(e, f). by definition.

(g) + left limits

Let's take some sequence $x_n \uparrow x$. Then

$$\{X < x\} = \bigcup_{n=1}^{\infty} \{X \le x_n\}$$

So Lemma $\Rightarrow$

$$P(X<x) = \lim_{n\to\infty} P(X \le x_n)$$

$$= \lim_{n\to\infty} F(x_n).$$

so $P(X<x) = \lim_{y \uparrow x} F(y)$.

so $P(X=x) = \mathbb{P}(X \le x) - P(X<x)$

$$= F(x) - \lim_{y\uparrow x} F(y).$$

Let's also finish proving that two r.v.s. have the same pmf iff they have the same cdf.

We already proved that if they have the same pmf they have the same cdf.

Suppose X and Y have the same cdf $F = F_X = F_Y$

Then

$$p_X(x) = P(X=x)$$

$$= F(x) - \lim_{y \uparrow x} F(y)$$

$$= P(Y=x)$$

$$= p_Y(x).$$

Independence of r.v.s

Defn. The r.v.s $X_1, X_2, \dots$ over a sample space Ω are called *independent* if $\forall$ (for all) $x_1, x_2, \dots \in \mathbb{R}$, the events

$$\underbrace{\{X_1 = x_1\}}_{\subseteq \Omega}, \underbrace{\{X_2 = x_2\}}_{\subseteq \Omega}, \dots$$

are independent.

Equivalently, if for any $A_1, A_2, \dots \subseteq \mathbb{R}$, the events

$$\{X_1 \in A_1\}, \{X_2 \in A_2\}, \dots$$

are independent.

Example $\Omega = \{H, T\}^2$

$P = \text{uniform}$

$$X_i(\omega_1, \omega_2) = 1\{\omega_i = H\}$$

Then X_1 & X_2 are independent.

because

$\{X_1 = 1\}$ & $\{X_2 = 1\}$ ind.

and $\{X_1 = 1\}$ & $\{X_2 = 0\}$ ind.

and $\{X_1 = 0\}$ & $\{X_2 = 1\}$ ind.

and $\{X_1 = 0\}$ & $\{X_2 = 0\}$ ind.

Proof. $P(X_1 = i_1 \text{ \& } X_2 = i_2)$ if $i_1, i_2 \in \{0,1\}$

$$= \sum_{\substack{\omega_1, \omega_2 \\ X_j(\omega) = i_j}} P(\omega_1, \omega_2) = \sum_{\substack{\omega \\ X_j(\omega) = i_j}} \frac{1}{4} = \frac{1}{4}.$$

Example Let $N \sim \text{Poisson}(\lambda)$.

Then run N iid $\text{Ber}(p)$

independent & identically distributed

trials. Let X be the number of successes and Y the number of failures. Then X & Y are independent. (even though $X+Y=N$ ... !)

(Poisson thinning).

Proof $P(X=x \mid N=n) = \binom{n}{x} p^x (1-p)^{n-x}$.

So $P(X=x, Y=y) = P(X=x, Y=y \mid N=x+y) P(N=x+y)$

$$= P(X=x \mid N=x+y) P(N=x+y) = \binom{x+y}{x} p^x (1-p)^y \cdot \frac{\lambda^{x+y}}{(x+y)!} e^{-\lambda}$$

$$= \underbrace{e^{-\lambda p} \frac{(\lambda p)^x}{x!}}_{P(X=x)} \underbrace{\frac{(\lambda(1-p))^y}{y!} e^{-\lambda(1-p)}}_{P(Y=y)}.$$

On the other hand,

$$P(X=x) = \sum_{n=x}^{\infty} P(X=x \mid N=n) P(N=n)$$

$$= \sum_{n=x}^{\infty} \binom{n}{x} p^x (1-p)^{n-x} \frac{\lambda^n}{n!} e^{-\lambda}$$

$$= p^x \lambda^x \frac{e^{-\lambda}}{x!} \sum_{n=x}^{\infty} \frac{(\lambda(1-p))^{n-x}}{(n-x)!}$$

$$= p^x \lambda^x \frac{e^{-\lambda}}{x!} \sum_{n=0}^{\infty} \frac{(\lambda(1-p))^n}{n!}$$

$$= p^x \lambda^x \frac{e^{-\lambda}}{x!} e^{\lambda(1-p)}$$

$$= \frac{(p\lambda)^x}{x!} e^{-\lambda p}.$$