Last update: Oct 16

List of theorems or properties:

Properties of a probability measure:

- $0 \leq P(A) \leq 1$
- $P(S) = 1$
- If $A \cap B = \emptyset$ then $P(A \cup B) = P(A) + P(B)$
- $P(A^c) = 1 - P(A)$
- $P(A) = P(A \cap B) + P(A \cap B^c)$
- $B \subseteq A$ then $P(A) = P(B) + P(A \cap B^c)$
- $B \subseteq A$ then $P(B) \leq P(A)$
- $A_1, A_2, \ldots$ a partition of S, then $P(B) = \sum_i P(B \cap A_i)$
- $P(A \cup B) = P(A) + P(B) + P(A \cap B)$
- $P(A_1 \cup A_2 \cup \ldots) \leq P(A_1) + P(A_2) + \ldots$

Combinatorics

- $n!$ is the number of permutations of a set of n elements (How many different arrangements)
- $\binom{n}{k} = \frac{n!}{k!(n-k!)} \rightarrow$ choose k elements from a set of n order doesn't matter
- What if order matters? $\binom{n}{k} \times k!$
- $\frac{n!}{(n-k!)}$ Number of ways to pick k items out of n where order does matter.
  - Is the same as $\binom{n}{k} \times k!$ Choosing without order and then calculating permutations of k items.

Conditional probability

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$ if $P(B) > 0$
- $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$
- $A_1, A_2, \ldots$ a partition of S, then $P(B) = \sum_i P(A_i)P(B|A_i) = \sum_i P(B \cap A_i)$
- $P(A|B) = \frac{P(A)}{P(B)} P(B|A)$

Independence

- A and B are independent if $P(A \cap B) = P(A)P(B)$
- $A_1, A_2, A_3, \ldots$ are independent if $P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2})\ldots P(A_{i_k})$ for any finite subcollection of events.

Continuity

- If $\{A_n\} \nearrow A$ then $\lim_{n\to\infty} P(A_n) = P(A)$
- If $\{A_n\} \searrow A$ then $\lim_{n\to\infty} P(A_n) = P(A)$

Random variables (r.v)

- A function $X : S \to \mathbb{R}$ is a random variable
- The distribution of a random variable is the collection of all of the probabilities of the variable being in every possible subset of $\mathbb{R}$.
- A r.v. is discrete if $\sum_{x\in\mathbb{R}} P(X = x) = 1$
- Probability function $p_X(k) := P(X = k)$

Discrete random variables

- $X \sim Bernoulli(\theta)$ if

$$p_X(k) = \theta^k (1-\theta)^{1-k} \, \mathbb{1}_{k\in\{0,1\}}$$

$$p_X(k) = \begin{cases} \theta & \text{if } k = 1 \\ 1-\theta & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

- $X \sim Binomial(n, \theta)$ if

$$p_X(k) = \binom{n}{k} \theta^k (1-\theta)^{n-k} \, \mathbb{1}_{k\in\{0,1,\ldots,n\}}$$

- $X \sim Geometric(\theta)$ if

$$p_X(k) = (1-\theta)^k \theta \, \mathbb{1}_{k\in\{0,1,\ldots\}}$$

  This one counts # of failures before the first success.
  there's a second parametrization of geometric counting the number of trials until first success. Check wikipedia!

$$p_Y(k) = (1-\theta)^{k-1} \theta \, \mathbb{1}_{k\in\{1,2,\ldots\}}$$

- $X \sim Poisson(\lambda)$ if

$$p_x(k) = e^{-\lambda}\frac{\lambda^k}{k!}\ \mathbb{1}_{k\in\{0,1,\dots\}}$$

Poisson approximation

If n is very large and θ very small then $Binomial(n,\theta)$ is well approximated by $Poisson(\lambda = n\theta)$

- $X \sim NegativeBinomial(r,\theta)$ if

  $$p_X(k) = \binom{r-1+k}{k}\theta^r(1-\theta)^k\ \mathbb{1}_{k\in\{0,1,\dots\}}$$

  This counts the number of misses before the r^{th} successwith θ as the probability of success.
- $X \sim Hypergeomtric(N,M,n)$ if

  $$p_X(k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\ \mathbb{1}_{\max\{0,n+M-N\}\leq k\leq \min\{n,M\}}$$

  Here we have a big population of size N composed b 2 sub-populations:
  - The population of interest (of size M)
  - and the rest of the population (of size $N-M$)

We're taking a sample of size n and X counts the number of elements, in that sample, that come from the population of interest.

Continuous random variables

- A random variable is continuous if $P(X=k)=0\ \forall k$
- A density function is "any" function $f:\mathbb{R}\to\mathbb{R}$ with:
  - $f(x)>0$
  - $\int_{-\infty}^{\infty} f(x)\ dx = 1$
- Then

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

- $X \sim Uniform(a,b)$ for $a<b$ if

  $$f(x) = \frac{1}{b-a}\ \mathbb{1}_{x\in(a,b)}$$

- $X \sim Exponential(\lambda)$ if

$$f(x) = \lambda e^{-\lambda x} \;\; \mathbb{1}_{x \in (0,\infty)}$$

- $X \sim Normal(\mu, \sigma^2)$ if

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

Cumulative distribution function

Any random variable X the cumulative Distribution Function (CDF) is the function J_X defined by:

$$F_X(x) = P(X \leq x) \; \forall x \in \mathbb{R}$$

- For X discrete

$$F_X(x) = \sum_{u \leq x} P(X = u)$$

- For X absolutely continuous

$$F_x(x) = \int_{-\infty}^{x} f_x(u) \; du$$

Properties:

- $0 \leq F_X(x) \leq 1 \; \forall x \in \mathbb{R}$
- If $x \leq y$ then $F_X(x) \leq F_X(y)$
- $\lim_{x \to -\infty} F_X(x) = 0$
- $\lim_{x \to +\infty} F_X(x) = 1$

Annex

Geometric series

Let a_0 be a constant and define $a_n = a_0 r^n$ as a geometric series

Define $S_n := \sum_{k=0}^{n} r^k$ the partial sums of the first $\{r^k\}_{k=0}^{n}$
note that $\sum_{k=0}^{n} a_k = \sum_{k=0}^{n} a_0 r^k = a_0 \sum_{k=0}^{n} r^k = a_0 S_n$

Now $S_n = r^0 + r^1 + \ldots r^n$ and $rS_n = r^1 + r^2 + \ldots r^{n+1}$, then

$S_n - rS_n = r^0 - r^{n+1}$
$S_n(1-r) = r^0 - r^{n+1}$
then

$$S_n = \frac{r^0 - r^{n+1}}{1-r} = \frac{1-r^{n+1}}{1-r} \qquad \text{if } r \neq 1$$

Now if $r \in (0,1)$

$$\sum_{k=0}^{\infty} r^k = \lim_{n\to\infty} \sum_{k=0}^{n} r^k = \lim_{n\to\infty} S_n = \frac{1}{1-r}$$

A little bit of topology

- Remember that both $\emptyset$ and $\mathbb{R}$ are both open and closed sets.
  Consider ONLY nested intervals
- Class of open intervals (a_n, b_n)
  Closed under arbitrary unions and finite intersections
  This means:
- $\bigcup_{n=1}^{\infty}(a_n, b_n)$ will be an open set (either open interval or union of open intervals)
- $\bigcap_{n=1}^{M}(a_n, b_n)$ will be an open set (either open interval or union of open intervals or $\emptyset$)
- $\bigcap_{n=1}^{\infty}(a_n, b_n)$ will be a closed set (closed interval, singleton, $\emptyset$)
- Class of closed intervals $[c_k, d_k]$
  Closed under arbitrary intersections and finite unions
- $\bigcap_{k=1}^{\infty}[c_k, d_k]$ will be a closed set (closed interval, singleton, $\emptyset$)
- $\bigcup_{k=1}^{\infty}[c_k, d_k]$ most likely will be open or have a side that is open $[c, d)$
- $\bigcap_{k=1}^{M}[c_k, d_k]$ will be a closed set

#probability
#ta