Cheat sheet

Last updated: March 3, 2020

Events, Probabilities

January 15th: Probability spaces

Def. Sample space S: set of all possible outcomes (ex: $S = \mathbb{R}$ for the temperature tomorrow).
Def. Event A: subset of A (ex: $A = [50, +\infty)$ for the temperature being too hot).
Logic connectors $\leftrightarrow$ set operations (ex: "A and B" $\leftrightarrow$ $A \cap B$; "not A" $\leftrightarrow$ $S \setminus A = A^{\complement}$).
Def. A and B disjoint: $A \cap B = \emptyset$ ($\leftrightarrow$ "A and B cannot occur simultaneously").
$A_1, A_2, \ldots, A_n, \ldots$ pairwise disjoint: A_n and A_m disjoint for $n \neq m$.
Def. Probability space $(S, \mathbb{P})$: sample space S and probability $\mathbb{P}(A)$ for every event $A \subset S$, with

- $0 \leq \mathbb{P}(A) \leq 1$,
- $\mathbb{P}\big(\bigcup_{n\geq 0} A_n\big) = \sum_{n\geq 0} \mathbb{P}(A_n)$ for $A_1, A_2, \ldots$ pairwise disjoints.

January 17th: Operations on events

Rk. For a finite sample set S, it suffices to know $\mathbb{P}(\{x\})$ for all $x \in S$.
Rk. Motto: the sample space is less important than the events.
Def. A is included in B: $A \subset B$ ($\leftrightarrow$ "A implies B").
Prop. $\mathbb{P}(A^{\complement}) = 1 - \mathbb{P}(A)$.
Prop. $\mathbb{P}(A \cup B) = 1 - \mathbb{P}(A^{\complement} \cap B^{\complement})$.
Prop. $\mathbb{P}(A \cap B) = 1 - \mathbb{P}(A^{\complement} \cup B^{\complement})$.
Prop. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
Prop. $\mathbb{P}(A) \leq \mathbb{P}(B)$ whenever $A \subset B$.

January 22nd: Operations on events / Conditional probability

Prop(Inclusion-Exclusion Principle)**.**

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$$

Def. $A_0, A_1, A_2, \ldots$ increasing: $A_0 \subset A_1 \subset A_2 \subset \cdots$.
$A_0, A_1, A_2, \ldots$ decreasing: $A_0 \supset A_1 \supset A_2 \supset \cdots$.
Prop. $\mathbb{P}(\bigcup_{n\geq 0} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$ for $(A_n)_{n\geq 0}$ increasing;

$\mathbb{P}(\bigcap_{n\geq 0} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)$ for $(A_n)_{n\geq 0}$ decreasing.

Counting

Rk. *Tuples* are ordered sequences, *sets* are unordered sequences.
They are denoted by $(a_1, \ldots, a_k)$ and $\{a_1, \ldots, a_k\}$ respectively.
Def. Conditional probability $\mathbb{P}(A|B)$ of A knowing B: probability of A happening considering only outcomes where B holds.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$$

Def. A and B independent: $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

January 24th: Product and uniform probability spaces

Def. Product set: $S_1 \times S_2 = \{(a, b),\ a \in S_1,\ b \in S_2\}$ ($\leftrightarrow$ experience S_1, then experience S_2).
Prop. For finite sets, $\#(A \times B) = (\#A) \times (\#B)$.
Def. Product probability space (for finite spaces): $S_1 \times S_2$ with $\mathbb{P}(\{(a, b)\}) = \mathbb{P}_1(\{a\})\mathbb{P}_2(\{b\})$ ($\leftrightarrow$ experience S_1, then experience S_2 independently).
Def. Uniform probability (finite set): all outcomes equally likely, i.e. $\mathbb{P}(\{a\}) = \mathbb{P}(\{b\})$.
Prop. For $(S, \mathbb{P})$ uniform, $\mathbb{P}(A) = \frac{\#A}{\#S}$ for every event $A \subset S$.
Thm. The product of two uniform probability spaces is again uniform.
Ex. Sampling with replacement: $U = \{1, \ldots, n\}$ an urn with n elements,
$S = U^k \leftrightarrow$ draw k elements with replacement.

January 27th: Sampling without replacement

Def. kth falling factorial: $(n)_k = n^{\underline{k}} = n \times (n-1) \times \cdots \times (n-k+1)$ (k terms).
Prop. If U has n elements, $\#\{k\text{-tuples of distinct elements of } U\} = (n)_k$.
Rk. The notation $U^{\underline{k}} = \{k\text{-tuples of distinct elements of } U\}$ is sometimes used.

January 29th: Permutations, combinations

Def. Permutations: $\mathfrak{S}(U) = \{\text{Permutations of } U\} = \{\text{Bijections } f : U \to U\}$.
Def. $\mathfrak{S}_n = \mathfrak{S}(\{1, \ldots, n\})$ ($\leftrightarrow$ orderings of $\{1, \ldots, n\}$ via $f(\text{element}) = \text{position}$).
Def. Factorial: $n! = n \times (n-1) \times \cdots \times 1$.
Prop. If U has n elements, $\#\mathfrak{S}(U) = n!$; e.g. $\#\mathfrak{S}_n = n!$.
Def. Binomial coefficient "n choose k": $\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}$.
Prop. If U has n elements, $\#\{\text{subsets of } U \text{ with } k \text{ elements}\} = \binom{n}{k}$.
Rk. The notation $\binom{U}{k} = \{\text{subsets of } U \text{ with } k \text{ elements}\}$ is sometimes used.

January 31st: Multinomials

Def. Multinomial coefficient: for $k_1 + \cdots + k_\ell = n$, $\binom{n}{k_1,\ldots,k_\ell} = \frac{n!}{k_1! \times \cdots \times k_\ell!}$.
Def. Partition of U: $S_1, \ldots, S_\ell \subset U$ such that $S_1 \cup \cdots \cup S_\ell = U$ and $S_i \cap S_j = \emptyset$.
Prop. If U has n elements, $\#\{\text{partitions } S_1, \ldots, S_\ell \text{ of } U \text{ such that } \#S_i = k_i\} = \binom{n}{k_1,\ldots,k_\ell}$.
Rk. The notation $\binom{U}{k_1,\ldots,k_\ell} = \{\text{partitions } S_1, \ldots, S_\ell \text{ of } U \text{ such that } \#S_i = k_i\}$ is sometimes used.

Independence, Conditioning

February 5th: Conditioning

Def. Conditional probability $\mathbb{P}(A|B)$ of A knowing B: probability of A happening considering only outcomes where B holds.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Prop. $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B)$.
Prop. $\mathbb{P}(-|B)$ is a probability function, hence $\mathbb{P}(A^{\complement}|B) = 1 - \mathbb{P}(A^{\complement}|B)$,
$\mathbb{P}(C \cup D|B) = \mathbb{P}(C|B) + \mathbb{P}(D|B) - \mathbb{P}(C \cap D|B)$, etc.
Prop. If $\mathbb{P}$ is uniform on S, then $\mathbb{P}(A|B) = \frac{\#(A \cap B)}{\#B}$.

February 7th: Independence 1

Def. Independence: A, B independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
Rk. If $\mathbb{P}(B) \neq 0$, it is equivalent to $\mathbb{P}(A|B) = \mathbb{P}(A)$, i.e. B has no influence on A.
Def. Geometric distribution: $N \sim \mathcal{G}eom(p)$ if $\mathbb{P}(N = n) = (1-p)^{n-1}p$ for $n \in \mathbb{N}^*$
($\leftrightarrow$ first success in a sequence of independent experiments with a probability p of success).
Def. Binomial distribution: $N \sim \mathcal{B}in(n,p)$ if $\mathbb{P}(N = k) = \binom{n}{k}(1-p)^{n-k}p^k$ for $0 \leq k \leq n$
($\leftrightarrow$ number of successes in a sequence of n independent experiments with a probability p of success).

February 10th: Independence 2, Law of total probability

Rk. When dealing with independent experiments, try to transform 'or' to 'and':
$\mathbb{P}(A \cup B) = 1 - \mathbb{P}(A^{\complement} \cap B^{\complement}) = 1 - \mathbb{P}(A^{\complement})\mathbb{P}(B^{\complement}) = 1 - (1 - \mathbb{P}(A))(1 - \mathbb{P}(B))$.
Prop. $\mathbb{P}(A \text{ before } B) = \frac{\mathbb{P}(A)}{\mathbb{P}(A)+\mathbb{P}(B)}$ (see the notes for a precise statement).
Def. Partition: $B_0, B_1, \ldots \subset S$ such that $S = \bigcup_{k=0}^{\infty} B_k$, $B_k \cap B_\ell = \emptyset$, $\mathbb{P}(B_k) \neq 0$.
Thm(law of total probability)**.** If $B_0, B_1, \ldots$ is a partition, then

$$\mathbb{P}(A) = \mathbb{P}(A|B_0)\mathbb{P}(B_0) + \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \cdots$$

February 12th: Bayes' law

Thm(Bayes' law)**.** If A and B have positive measure, then $\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)}$.
If A has positive measure and $B_0, B_1, \ldots$ is a partition, then

$$\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A|B_0)\mathbb{P}(B_0) + \mathbb{P}(A|B_1)\mathbb{P}(B_1) + \cdots}$$

Discrete random variables

Def. Random variable: function $X : S \to \mathbb{R}$, with $(S, \mathbb{P})$ a probability space
($\leftrightarrow$ random number deduced from the outcome $x \in S$ of an experiment).
Def. Discrete random variable: random variable with values in a countable set $\{x_0, x_1, \ldots\}$.
Def. Uniform distribution: $N \sim \mathcal{U}nif(S)$ if $\mathbb{P}(N = x) = \frac{1}{\#S}$ for $x \in S$.

February 14th: Usual distributions 1, cumulative distribution function 1

Def. Probability mass function: for N discrete, $x \mapsto \mathbb{P}(N = x)$.
Def. Bernoulli distribution: $N \sim \mathcal{B}er(p)$ if $\mathbb{P}(N = 1) = 1 - \mathbb{P}(N = 0) = p$.
Def. Geometric distribution: $N \sim \mathcal{G}eom(p)$ if $\mathbb{P}(N = n) = (1 - p)^{n-1} p$ for $n \in \mathbb{N}^*$.
Def. Binomial distribution: $N \sim \mathcal{B}in(n, p)$ if $\mathbb{P}(N = k) = \binom{n}{k}(1 - p)^{n-k} p^k$ for $0 \leq k \leq n$.
Def. Cumulative distribution function of N: $F_N : x \mapsto \mathbb{P}(N \leq x)$.
Prop. F_N is non-descreasing, tends to zero (resp. one) as $-\infty$ (resp. $+\infty$), and is right continuous.

February 16th: Usual distributions 2

Def. Negative binomial distribution: $N \sim \mathcal{N}eg\mathcal{B}in(k, p)$ if $\mathbb{P}(N = n) = \binom{n-1}{k-1}(1 - p)^{n-k} p^k$
for $n \in \mathbb{N}, n \geq k$
($\leftrightarrow$ kth success in a sequence of independent experiments with a probability p of success).
Def. Hypergeometric distribution: $N \sim \mathcal{H}yper\mathcal{G}eom(w, b, n)$ if $\mathbb{P}(N = k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$
for $k \in \mathbb{N}, k \geq n - b, k \leq w, k \leq n$.
($\leftrightarrow$ number of whites out of n balls sampled without replacement from a pool of w white and b black balls).

February 19th: Cumulative distribution function, quantiles

Prop. The cumulative distribution function of a discrete random variable N consists in horizontal steps satisfying the conditions of monotonicity, convergence and right-continuity.
Prop. The x-coordinate x of a jump of F_N corresponds to a possible value of N.
The height of the jump is $\mathbb{P}(N = x)$.
Def. Quartile: the p-quantile of N is the smallest x such that $\mathbb{P}(N \leq x) \geq p$.
Def. Percentile: the q-percentile of N is its $\frac{q}{100}$-quantile.
Def. Quartile: the lower, or first quartile, is the 1/4-quantile;
the upper, or third quartile, is the 3/4-quantile.
Def. Median: the median, or second quartile, is the 1/2-quantile.

February 26th: Poisson distributions

Def. Poisson distribution $\mathcal{P}oi(\lambda)$: $N \sim \mathcal{P}oi(\lambda)$ if $\mathbb{P}(N = k) = \mathrm{e}^{-\lambda} \cdot \frac{\lambda^k}{k!}$.
($\leftrightarrow$ number of successes for an experiment conducted continuously, for an average of λ successes during that time period).

February 28th: Independence of random variables

Def. Independence: N, M are independent random variables
if $\mathbb{P}(N \in A \cap M \in B) = \mathbb{P}(N \in A) \cdot \mathbb{P}(M \in B)$ for all A, B.

Prop. If $N \sim \mathcal{B}in(n, p)$, $M \sim \mathcal{B}in(m, p)$ are independent, then $N + M \sim \mathcal{B}in(n + m, p)$.
If $N \sim \mathcal{N}eg\mathcal{B}in(k, p)$, $M \sim \mathcal{N}eg\mathcal{B}in(\ell, p)$ are independent, then $N + M \sim \mathcal{N}eg\mathcal{B}in(k + \ell, p)$.
If $N \sim \mathcal{P}oi(\lambda)$, $M \sim \mathcal{P}oi(\mu)$ are independent, then $N + M \sim \mathcal{N}eg\mathcal{B}in(\lambda + \mu)$.

Rk. If $N_1, \ldots, N_n$ are independent with distribution $\mathcal{B}er(p)$, then $N_1 + \cdots + N_n \sim \mathcal{B}in(n, p)$.
If $N_1, \ldots, N_k$ are independent with distribution $\mathcal{G}eom(p)$, then $N_1 + \cdots + N_k \sim \mathcal{N}eg\mathcal{B}in(k, p)$.

Statistical invariants of discrete random variables

March 2nd: Expectation

Def. Expectation: For N a discrete random variable with values in $\{x_0, x_1, x_2, \ldots\}$,
$\mathbb{E}[N] = \sum_{k=0}^{\infty} x_k \cdot \mathbb{P}(N = x_k)$.
If N is a non-negative integer, $\mathbb{E}[N] = \sum_{k=0}^{\infty} k \cdot \mathbb{P}(N = k) = \sum_{k=1}^{\infty} k \cdot \mathbb{P}(N = k)$.

Prop. For $N \sim \mathcal{B}er(p)$, $\mathbb{E}[N] = p$.
For $N \sim \mathcal{U}nif(\{a, \cdots, b\})$, $\mathbb{E}[N] = \frac{a+b}{2}$.
For $N \sim \mathcal{G}eom(p)$, $\mathbb{E}[N] = 1/p$.
For $N \sim \mathcal{P}oi(\lambda)$, $\mathbb{E}[N] = \lambda$.

Prop. For X, Y random variables, and a a deterministic real number,

$$\mathbb{E}[aN] = a\mathbb{E}[N], \qquad \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y], \qquad \mathbb{E}[1] = 1.$$

Prop. For $N \sim \mathcal{B}in(n, p)$, $\mathbb{E}[N] = np$.
For $N \sim \mathcal{N}eg\mathcal{B}in(k, p)$, $\mathbb{E}[N] = k/p$.
For $N \sim \mathcal{H}yper\mathcal{G}eom(w, b, n)$, $\mathbb{E}[N] = \frac{nw}{w+b}$.