ECE 3100 Probability Cheat Sheet (Ch. 1–2.3)

Sets (1.1)

Set: collection of elements. $x \in S$: x in S. **Empty set:** $\emptyset$. $S \subset T$: every element of S in T. **Universal set** Ω. **Complement:** $S^c = \{x \in \Omega \mid x \notin S\}$. **Union:** $S \cup T$. **Intersection:** $S \cap T$. **Disjoint:** $S \cap T = \emptyset$. **Partition** of S: disjoint sets whose union is S.

Set Identities

$(S^c)^c = S$, $S \cup \Omega = \Omega$, $S \cap \Omega = S$, $S \cup S^c = \Omega$, $S \cap S^c = \emptyset$, $S \cup \emptyset = S$, $S \cap \emptyset = \emptyset$

De Morgan's: $(S \cup T)^c = S^c \cap T^c$, $(S \cap T)^c = S^c \cup T^c$. Generalized: $(\bigcup_n S_n)^c = \bigcap_n S_n^c$, $(\bigcap_n S_n)^c = \bigcup_n S_n^c$

Distributive: $S \cap (T \cup U) = (S \cap T) \cup (S \cap U)$; $S \cup (T \cap U) = (S \cup T) \cap (S \cup U)$

Decomposition: $A^c = (A^c \cap B) \cup (A^c \cap B^c)$; $(A \cap B)^c = (A^c \cap B) \cup (A^c \cap B^c) \cup (A \cap B^c)$

Probabilistic Models (1.2)

An **experiment** produces one outcome from **sample space** Ω. **Event**: subset of Ω. Outcomes must be mutually exclusive and collectively exhaustive.

Probability Axioms

1. **Nonnegativity:** $P(A) \geq 0$ for every event A
2. **Additivity:** If $A \cap B = \emptyset$: $P(A \cup B) = P(A) + P(B)$. For countably many disjoint $A_1, A_2, \ldots$: $P(\bigcup_i A_i) = \sum_i P(A_i)$
3. **Normalization:** $P(\Omega) = 1$

Consequences of Axioms

$P(\emptyset) = 0$; $P(A^c) = 1 - P(A)$; $0 \leq P(A) \leq 1$

If $A \subset B$, then $P(A) \leq P(B)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \cup B) \leq P(A) + P(B)$ (Union bound)

$P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$

Inclusion-Exclusion: $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$

Bonferroni: $P(A \cap B) \geq P(A) + P(B) - 1$; general: $P(\bigcap_{i=1}^n A_i) \geq \sum_{i=1}^n P(A_i) - (n-1)$

Exactly one of A, B**:** $P(A \cap B^c) + P(A^c \cap B) = P(A) + P(B) - 2P(A \cap B)$

Discrete Probability Law: If $\Omega = \{s_1, \ldots, s_n\}$ is finite: $P(\{s_1, \ldots, s_n\}) = P(s_1) + \cdots + P(s_n)$

Discrete Uniform: all outcomes equally likely $\Rightarrow P(A) = \frac{|A|}{|\Omega|}$

Conditional Probability (1.3)

$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$, $P(B) > 0$. $\Rightarrow P(A \cap B) = P(B)\,P(A \mid B) = P(A)\,P(B \mid A)$

Cond. prob. is a valid prob. law—all axiom-derived properties hold. If equally likely: $P(A \mid B) = \frac{|A \cap B|}{|B|}$. $P(A^c \mid B) = 1 - P(A \mid B)$

If $A_1 \cap A_2 = \emptyset$: $P(A_1 \cup A_2 \mid B) = P(A_1 \mid B) + P(A_2 \mid B)$

Multiplication Rule

$P(\bigcap_{i=1}^n A_i) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap \cdots \cap A_{n-1})$

Tree method: conditional probs on branches; leaf prob = product along path; event prob = sum of its leaves.

Total Probability & Bayes' Rule (1.4)

Let $A_1, \ldots, A_n$ be disjoint events forming a **partition** of Ω with $P(A_i) > 0$ for all i.

Total Probability Theorem

$P(B) = \sum_{i=1}^n P(A_i)\,P(B \mid A_i)$ (weighted avg of cond. probs by scenario)

Bayes' Rule

$P(A_i \mid B) = \frac{P(A_i)\,P(B \mid A_i)}{\sum_{j=1}^n P(A_j)\,P(B \mid A_j)} = \frac{P(A_i)\,P(B \mid A_i)}{P(B)}$ ($P(A_i)$: prior; $P(A_i \mid B)$: posterior)

Two-event: $P(A \mid B) = \frac{P(A)\,P(B \mid A)}{P(A)\,P(B \mid A) + P(A^c)\,P(B \mid A^c)}$

False-Positive Puzzle: An accurate test can have low positive predictive value if the prior probability (prevalence) is very small.

Independence (1.5)

A and B are **independent** if and only if: $P(A \cap B) = P(A)\,P(B)$

Equivalently (if $P(B) > 0$): $P(A \mid B) = P(A)$

Key facts about independence:

- A, B indep. $\Rightarrow$ A, B^c also indep.; A^c, B also indep.; A^c, B^c also indep.
- **Disjoint $\neq$ independent!** If $A \cap B = \emptyset$ and $P(A) > 0, P(B) > 0$, then A, B are **NOT independent** since $P(A \cap B) = 0 \neq P(A)P(B)$. Disjoint events with positive probability are *always* dependent.
- A and its complement A^c are NOT independent (unless $P(A) = 0$ or 1).
- If $P(A) = 0$ or $P(A) = 1$, then A is independent of every event.
- Independence is a *symmetric* property: A indep. of $B \Leftrightarrow B$ indep. of A.

Disjoint vs. Independent – Key Comparison

Disjoint ($A \cap B = \emptyset$)	**Independent**
$P(A \cap B) = 0$	$P(A \cap B) = P(A)P(B)$
$P(A \cup B) = P(A) + P(B)$	$P(A \cup B) = P(A) + P(B) - P(A)P(B)$
$P(A \mid B) = 0$ (if $P(B) > 0$)	$P(A \mid B) = P(A)$
Cannot both occur	Can both occur
Knowing B occurred $\Rightarrow$ A did not	Knowing B tells nothing about A
$P(A), P(B) > 0 \Rightarrow$ NOT indep.	Can be disjoint only if $P(A) = 0$ or $P(B) = 0$

Conditional Independence given C ($P(C) > 0$): $P(A \cap B \mid C) = P(A \mid C)\,P(B \mid C)$; equiv. $P(A \mid B \cap C) = P(A \mid C)$.

Indep. $\not\Rightarrow$ cond. indep., and vice versa.

Independence of Multiple Events

$A_1, \ldots, A_n$ **independent** if for *every* subset $S \subseteq \{1, \ldots, n\}$: $P(\bigcap_{i \in S} A_i) = \prod_{i \in S} P(A_i)$. For 3 events, need all four: $P(A_i \cap A_j) = P(A_i)P(A_j)$ for each pair AND $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$. Pairwise indep. $\not\Rightarrow$ full indep.; triple condition alone $\not\Rightarrow$ pairwise. n events: need $2^n - n - 1$ conditions.

Bernoulli Trials

n independent tosses, $P(\text{head}) = p$: $P(k \text{ heads}) = \binom{n}{k} p^k (1-p)^{n-k}$

Reliability

Series (all work): $P = p_1 p_2 \cdots p_m$. **Parallel** (any works): $P = 1 - (1-p_1) \cdots (1-p_m)$.

Counting (1.6)

Counting Principle

r-stage process with n_i choices at stage i (regardless of prior choices): total outcomes $= n_1 \cdot n_2 \cdots n_r$

Permutations

k**-permutations** (ordered, k from n): $\frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1)$

Combinations (Binomial Coefficients)

Choose k from n (unordered): $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Properties: $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{k} = \binom{n}{n-k}$

Pascal's Rule: $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

$\sum_{k=0}^{n} \binom{n}{k} = 2^n$ (total subsets of n-element set)
Binomial formula: $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$
$k\binom{n}{k} = n\binom{n-1}{k-1}$; $\sum_{k=0}^{n} k\binom{n}{k} = n \cdot 2^{n-1}$

Multinomial Coefficient (Partitions)

Partition n objects into r groups of sizes $n_1, \ldots, n_r$ (where $\sum n_i = n$): $\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1!\, n_2! \cdots n_r!}$
Useful for anagrams: TATTOO $= \frac{6!}{3!\,1!\,2!} = 60$.

Sampling Summary

- **Ordered, with replacement:** n^k sequences
- **Ordered, without replacement:** $\frac{n!}{(n-k)!}$ (k-permutations)
- **Unordered, without replacement:** $\binom{n}{k}$ (combinations)

Discrete Random Variables (2.1–2.3)

A **random variable** (RV) X: real-valued function of outcome. **Discrete**: finitely/countably many values.

Probability Mass Function (PMF) (2.2)

$p_X(x) = P(X = x)$
Properties: $p_X(x) \geq 0$ for all x; $\sum_x p_X(x) = 1$
$P(X \in S) = \sum_{x \in S} p_X(x)$. To compute: for each x, collect all outcomes giving $X = x$, sum their probs.

Common Discrete RVs

Bernoulli(p): $X \in \{0, 1\}$, $p_X(1) = p$, $p_X(0) = 1-p$. $E[X] = p$, $\text{var}(X) = p(1-p)$.
Binomial(n, p): X = #successes in n indep. trials. $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, \ldots, n$. $E[X] = np$, $\text{var}(X) = np(1-p)$.
Geometric(p): X = #trials until 1st success. $p_X(k) = (1-p)^{k-1} p$, $k = 1, 2, \ldots$ $E[X] = \frac{1}{p}$, $\text{var}(X) = \frac{1-p}{p^2}$. $P(X > k) = (1-p)^k$. Memoryless: $P(X > m+n \mid X > m) = P(X > n)$.
Poisson(λ): $p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k = 0, 1, \ldots$ $E[X] = \lambda$, $\text{var}(X) = \lambda$. Approx. Binom(n, p) when n large, p small, $\lambda = np$.
Discrete Uniform on $\{a, \ldots, b\}$: $p_X(k) = \frac{1}{b-a+1}$. $E[X] = \frac{a+b}{2}$, $\text{var}(X) = \frac{(b-a)(b-a+2)}{12}$. Die: $E[X] = 3.5$, $\text{var}(X) = 35/12$.

Functions of Random Variables (2.3)

If $Y = g(X)$, then Y is also a discrete RV: $p_Y(y) = \sum_{\{x: g(x) = y\}} p_X(x)$. E.g., $Y = |X|$: $p_Y(y) = p_X(y) + p_X(-y)$ for $y > 0$.

Expectation, Mean & Variance (2.4)

Expected Value (Mean)

$E[X] = \sum_x x\, p_X(x)$ (center of gravity of PMF)
Expected Value Rule: $E[g(X)] = \sum_x g(x)\, p_X(x)$ (no need to find PMF of $g(X)$!)
n-th moment: $E[X^n] = \sum_x x^n p_X(x)$. **Warning:** $E[g(X)] \neq g(E[X])$ in general (only if g linear).

Variance & Standard Deviation

$\text{var}(X) = E[(X - E[X])^2] = \sum_x (x - E[X])^2 p_X(x)$
Shortcut: $\text{var}(X) = E[X^2] - (E[X])^2$. $\sigma_X = \sqrt{\text{var}(X)}$. $\text{var}(X) \geq 0$; $= 0$ iff X constant w.p. 1.

Linear Functions

$Y = aX + b$: $E[Y] = aE[X] + b$, $\text{var}(Y) = a^2\, \text{var}(X)$. Adding constant shifts mean, doesn't change variance.

Problem-Solving Strategies

Event Expressions

$P(\text{at least one of } A, B, C) = 1 - P(A^c \cap B^c \cap C^c)$
$P(\text{at most one of } A, B, C)$:
$= P(A^c \cap B^c \cap C^c) + P(A \cap B^c \cap C^c) + P(A^c \cap B \cap C^c) + P(A^c \cap B^c \cap C)$
$P(\text{at least 2 of } A, B, C)$: event $(A \cap B) \cup (A \cap C) \cup (B \cap C)$
$P(\text{exactly one of } A, B, C) = P(A) + P(B) + P(C) - 2P(A \cap B) - 2P(A \cap C) - 2P(B \cap C) + 3P(A \cap B \cap C)$
$P(\text{exactly 2 of } A, B, C) = P(A \cap B) + P(A \cap C) + P(B \cap C) - 3P(A \cap B \cap C)$

Express using set operations:

- "A occurs but B doesn't": $A \cap B^c$
- "Neither A nor B": $A^c \cap B^c = (A \cup B)^c$
- "Either A or B but not both" (XOR): $(A \cap B^c) \cup (A^c \cap B)$
- "A or, if not, then not B either": $A \cup B^c$

Solution Methods

Counting method: when Ω finite, outcomes equally likely. Count $|A|$ and $|\Omega|$, use $P(A) = |A|/|\Omega|$.
Sequential/tree method: for experiments with sequential stages. Record conditional probabilities on branches. Use multiplication rule for paths. Use addition for combining paths corresponding to an event.
Divide and conquer: partition Ω into scenarios $A_1, \ldots, A_n$. Use total probability theorem to find $P(B) = \sum P(A_i) P(B \mid A_i)$.
Inference/Bayes': observe "effect" B, want "cause" A_i. Apply Bayes' rule.
Complement method: $P(A) = 1 - P(A^c)$. Often easier to compute $P(A^c)$.

Useful Formulas & Identities

$\sum_{n=0}^{\infty} \gamma^n = \frac{1}{1-\gamma}$, $|\gamma| < 1$ (geometric series)
$\sum_{n=1}^{\infty} n\gamma^{n-1} = \frac{1}{(1-\gamma)^2}$, $|\gamma| < 1$
$\lim_{n\to\infty} \left(1 + \frac{\alpha}{n}\right)^n = e^{\alpha}$
$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$; $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$
$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$ (Stirling's approximation)
$0! = 1$; $\binom{n}{k} = 0$ if $k < 0$ or $k > n$

Derangements (permutations with no fixed points): $D_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} \approx \frac{n!}{e}$
Prob. that all n people draw own name: $\frac{1}{n!}$
Prob. first m draw own names: $\frac{(n-m)!}{n!} = \frac{1}{n(n-1)\cdots(n-m+1)}$
Hypergeometric distribution: N items, K "successes." Draw k without replacement: $P(X = x) = \frac{\binom{K}{x}\binom{N-K}{k-x}}{\binom{N}{k}}$
Gambler's Ruin: Start with \$$k$, win \$1 w.p. p, lose \$1 w.p. $q = 1-p$ each round, stop at \$0 or \$$n$. $P(\text{reach } n) = \begin{cases} \frac{1-(q/p)^k}{1-(q/p)^n} & p \neq q \\ k/n & p = q = 1/2 \end{cases}$
Best-of-$(2m-1)$ series: Team wins game w.p. p (indep.). $P(\text{win series}) = \sum_{k=m}^{2m-1} \binom{2m-1}{k} p^k (1-p)^{2m-1-k}$. Alternatively, team wins in exactly $m + j$ games ($j = 0, \ldots, m-1$): $\binom{m-1+j}{j} p^m (1-p)^j$.
Repetition coding: Send bit n times, decode by majority rule. Prob. of correct decoding $= \sum_{k=\lceil n/2 \rceil}^{n} \binom{n}{k} (1-p_e)^k p_e^{n-k}$ where p_e is bit-flip prob.
Sampling w/o replacement quality control: N items, M defective. Test K. Prob. no defectives found: $\frac{\binom{N-M}{K}}{\binom{N}{K}}$.