ECE 3100 Probability Cheat Sheet (Bertsekas & Tsitsiklis, Ch. 1–2.3)

1. Sets & Sample Spaces (Sec. 1.1)

Experiment: a procedure that produces exactly one out of several possible **outcomes**. **Sample space** Ω: set of *all* possible outcomes. **Event**: a subset $A \subseteq \Omega$. An event *occurs* if the outcome $\omega \in A$.

Set operations: $A \cup B$ (union/"or"), $A \cap B$ (intersection/"and"), $A^c = \Omega \setminus A$ (complement/"not A").

De Morgan's Laws: $(A \cup B)^c = A^c \cap B^c$; $(A \cap B)^c = A^c \cup B^c$. Generalized: $\big(\bigcup_i A_i\big)^c = \bigcap_i A_i^c$; $\big(\bigcap_i A_i\big)^c = \bigcup_i A_i^c$.

Distributive: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$; $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Partition: $A_1, \dots, A_n$ partition Ω if $A_i \cap A_j = \varnothing$ for $i \neq j$ and $\bigcup_i A_i = \Omega$.

Disjoint (Mutually Exclusive): $A \cap B = \varnothing$; A and B cannot both occur.

Subset: $A \subseteq B$ means every outcome in A is also in B; $P(A) \le P(B)$.

Complement partition: $A^c = (A^c \cap B) \cup (A^c \cap B^c)$; $(A \cap B)^c = (A^c \cap B) \cup (A^c \cap B^c) \cup (A \cap B^c)$.

Expressing events: "At least two of A, B, C": $(A \cap B) \cup (A \cap C) \cup (B \cap C)$. "Exactly one of A, B, C": $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$. "At most one of A, B, C": $(A \cap B)^c \cap (A \cap C)^c \cap (B \cap C)^c$. "$A$ or, if not, not B": $A \cup B^c$.

2. Probability Axioms (Sec. 1.2)

(i) **Nonnegativity:** $P(A) \ge 0$ for all A.
(ii) **Normalization:** $P(\Omega) = 1$.
(iii) **(Countable) Additivity:** If $A_1, A_2, \dots$ pairwise disjoint, $P\big(\bigcup_i A_i\big) = \sum_i P(A_i)$.

Key Consequences

$P(\varnothing) = 0$. $P(A^c) = 1 - P(A)$. $0 \le P(A) \le 1$.

Inclusion–Exclusion (2): $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Inclusion–Exclusion (3): $P(A \cup B \cup C) = \sum P(\cdot) - \sum P(\cdot \cap \cdot) + P(A \cap B \cap C)$.

Complement rule: $P(\text{at least one of } A, B) = 1 - P(A^c \cap B^c)$.

Union bound (Boole): $P(A \cup B) \le P(A) + P(B)$; equality iff disjoint. $P\big(\bigcup_i A_i\big) \le \sum_i P(A_i)$.

Difference: $P(A \setminus B) = P(A \cap B^c) = P(A) - P(A \cap B)$.

Bounds on $P(A \cap B)$: $\max(0,\ P(A) + P(B) - 1) \le P(A \cap B) \le \min(P(A), P(B))$. $\max(P(A), P(B)) \le P(A \cup B) \le \min(1,\ P(A) + P(B))$.

Discrete Uniform Law

If Ω finite with $|\Omega| = n$ equally likely outcomes: $P(A) = |A|/|\Omega| = (\#\text{ favorable})/(\#\text{ total})$.

Continuous Uniform Models

Uniform on interval $[a, b]$: $P([c, d]) = (d - c)/(b - a)$ for $a \le c \le d \le b$.

Uniform on region $S \subset \mathbb{R}^2$: $P(A) = \text{Area}(A \cap S)/\text{Area}(S)$.

Manhattan distance: $|x| + |y|$. Point uniform on $[0, 1]^2$: $P(x + y \le a)$: if $0 \le a \le 1$, $= a^2/2$; if $1 < a \le 2$, $= 1 - (2 - a)^2/2$.

Meeting problem: Two arrivals uniform on $[0, T]$. $P(|X - Y| \le w) = 1 - (1 - w/T)^2$ for $0 \le w \le T$. One arrives first but other is late by $> w$: geometric region, $P = \frac{(T-w)^2}{2T^2}$ (each person).

3. Conditional Probability (Sec. 1.3)

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

$P(\cdot \mid B)$ is itself a valid probability law on Ω (satisfies all three axioms). So conditional versions of all rules hold, e.g. $P(A_1 \cup A_2 \mid B) = P(A_1 \mid B) + P(A_2 \mid B)$ when $A_1 \cap A_2 = \varnothing$.

Multiplication rule: $P(A \cap B) = P(B)\, P(A \mid B) = P(A)\, P(B \mid A)$.

Chain rule: $P\big(\bigcap_{i=1}^n A_i\big) = P(A_1) \prod_{k=2}^n P\big(A_k \mid \bigcap_{j=1}^{k-1} A_j\big)$.

Example (two coins): $P(\text{both H} \mid \text{first H}) = p$ (just need second H). $P(\text{both H} \mid \text{at least one H}) = p^2/(1 - (1 - p)^2) = p^2/(2p - p^2)$; this is $\le p$ for $p \in (0, 1)$.

4. Total Probability Theorem (Sec. 1.3)

If $A_1, \dots, A_n$ partition Ω with $P(A_i) > 0$: $P(B) = \sum_{i=1}^n P(A_i)\, P(B \mid A_i)$.

Use: break a complex event B into simpler conditional scenarios. E.g. radar: $P(\text{blip}) = P(\text{blip} \mid \text{alien}) P(\text{alien}) + P(\text{blip} \mid \text{no alien}) P(\text{no alien})$.

5. Bayes' Rule (Sec. 1.4)

$$P(A_i \mid B) = \frac{P(B \mid A_i)\, P(A_i)}{\sum_{j=1}^n P(B \mid A_j)\, P(A_j)}.$$

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

Terminology: Prior $P(A_i)$—initial belief. **Likelihood** $P(B \mid A_i)$—how likely evidence is under each hypothesis. **Posterior** $P(A_i \mid B)$—updated belief after evidence.

Sequential/iterative Bayes: After observing B_1, posterior $P(A_i \mid B_1)$ becomes the new prior; observe B_2 and apply Bayes again with $P(A_i \mid B_1)$ as prior.

False positive/negative: $P(\text{false alarm}) = P(\text{detect} \mid \text{absent})\, P(\text{absent})$. $P(\text{miss}) = P(\text{no detect} \mid \text{present})\, P(\text{present})$.

Monty Hall / Prisoner: Posterior depends on the guard's/host's randomization strategy when the player's situation allows multiple reveals. Often the "naive" conditional reasoning is wrong.

6. Independence (Sec. 1.5)

Two Events

A and B are **independent** iff $P(A \cap B) = P(A)\, P(B)$.

Equivalent (when defined): $P(A \mid B) = P(A)$, $P(B \mid A) = P(B)$.

If $A \perp\!\!\!\perp B$ then: $A \perp\!\!\!\perp B^c$, $A^c \perp\!\!\!\perp B$, $A^c \perp\!\!\!\perp B^c$.

Independence vs. Disjointness — Key Comparison

Disjoint ($A \cap B = \varnothing$)	**Independent** ($P(A \cap B) = P(A)P(B)$)
$P(A \cap B) = 0$	$P(A \cap B) = P(A)P(B)$
$P(A \cup B) = P(A) + P(B)$	$P(A \cup B) = P(A) + P(B) - P(A)P(B)$
$P(A \mid B) = 0$ (if $P(B) > 0$)	$P(A \mid B) = P(A)$
Knowing B occurred $\Rightarrow$ A did not	Knowing B occurred gives *no info* about A

Critical fact: If $P(A) > 0$ and $P(B) > 0$, disjoint events are **never** independent (since $0 \neq P(A)P(B)$). Disjoint events are *maximally dependent*—occurrence of one *rules out* the other.

Exception: If $P(A) = 0$ (or $P(B) = 0$), then A and B can be both disjoint and independent.

Positive/Negative Association

$P(A \mid B) > P(A) \Leftrightarrow P(B \mid A) > P(B)$ (symmetric). This means A, B are *not* independent and *not* disjoint (when both have positive prob). If $P(A \mid B) > P(A)$, can A, B be independent? **No.** If $P(A \mid B) > P(A)$, can A, B be disjoint? **No** (would need $P(A \mid B) = 0 < P(A)$).

Multiple Events & Mutual Independence

$A_1, \dots, A_n$ **mutually independent** iff for *every* subset $S \subseteq \{1, \dots, n\}$ with $|S| \ge 2$: $P\big(\bigcap_{i \in S} A_i\big) = \prod_{i \in S} P(A_i)$.

For 3 events: need $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$, $P(B \cap C) = P(B)P(C)$, *and* $P(A \cap B \cap C) = P(A)P(B)P(C)$. Pairwise $\not\Rightarrow$ mutual.

Key identity (complements): If $A_1, \dots, A_n$ mutually independent: $P(A_1^c \cap \cdots \cap A_n^c) = \prod_{i=1}^n (1 - P(A_i))$; $P(A_1^c \cup \cdots \cup A_n^c) = 1 - \prod_{i=1}^n P(A_i)$.

Independent of itself: $P(A) = P(A)^2 \Rightarrow P(A) \in \{0, 1\}$.

Independent trials: Coin flips, die rolls, transmissions—each trial's outcome does not affect others. Product rule applies: $P(\text{seq}) = \prod P(\text{each})$.

7. Counting (Sec. 1.6)

Multiplication principle: r stages with $n_1, n_2, \dots, n_r$ choices $\Rightarrow \prod n_i$ total.

Permutations (all n): $n!$; $0! = 1$. **k-permutations:** $n!/(n-k)!$ (ordered subsets of size k).

Combinations: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ (unordered subsets). $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{k} = \binom{n}{n-k}$, $\sum_{k=0}^n \binom{n}{k} = 2^n$.

Multinomial: $\frac{n!}{n_1! n_2! \cdots n_r!}$ ways to partition n into groups of sizes $n_1 + \cdots + n_r = n$.

Sampling summary:

	Ordered	Unordered
With replacement	n^k	$\binom{n+k-1}{k}$
Without replacement	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

Hypergeometric: N items, M defective, draw K w/o replacement. $P(\text{exactly } j \text{ defective}) = \frac{\binom{M}{j}\binom{N-M}{K-j}}{\binom{N}{K}}$, $0 \le j \le \min(K, M)$. $P(\text{no defective}) = \frac{\binom{N-M}{K}}{\binom{N}{K}}$; this *decreases* as M increases (for fixed K).

Sum formula: $\sum_{m=1}^n m = \frac{n(n+1)}{2}$.

Quality control: Reject batch if ≥ 1 defective in sample. $P(\text{reject} \mid M \text{ defective}) = 1 - \frac{\binom{N-M}{K}}{\binom{N}{K}}$.

Random assignment: n items to n people: $n!$ arrangements. $P(\text{all match}) = 1/n!$; $P(\text{first } m \text{ match}) = (n-m)!/n!$. $P(\text{first } m \text{ get names of last } m) = \binom{m}{m} \cdot m! \cdot (n-m)!/n! = m!(n-m)!/n!$.

Rooks on chessboard: 8 rooks on distinct squares of 8×8, all safe (no shared row/col): $P = \frac{8! \cdot \binom{8}{8} \cdot 8!}{\binom{64}{8} \cdot 8!} = \frac{8!}{\binom{64}{8}}$.

8. Discrete Random Variables (Sec. 2.1–2.3)

Random variable (r.v.): function $X : \Omega \to \mathbb{R}$. **Discrete:** range is finite or countably infinite.

8.1 PMF (Probability Mass Function) (Sec. 2.1)

$p_X(x) = P(X = x)$. $p_X(x) \ge 0$; $\sum_x p_X(x) = 1$. $P(X \in S) = \sum_{x \in S} p_X(x)$.

8.2 Common Discrete Distributions

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. Bernoulli trials. $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, \dots, n$. $E[X] = np$, $\text{Var}(X) = np(1-p)$.

Geometric(p): $X = \#$ trials until first success. $p_X(k) = (1-p)^{k-1} p$, $k = 1, 2, \dots$ $E[X] = 1/p$, $\text{Var}(X) = (1-p)/p^2$. $P(X > k) = (1-p)^k$. **Memoryless:** $P(X > m + n \mid X > m) = P(X > n)$.

Negative Binomial (Pascal): $X = \#$ trials until r-th success. $p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$, $k = r, r+1, \dots$ $E[X] = r/p$, $\text{Var}(X) = r(1-p)/p^2$.

Discrete Uniform on $\{a, \dots, b\}$: $p_X(k) = 1/(b - a + 1)$. $E[X] = (a+b)/2$. $\text{Var}(X) = (b-a)(b-a+2)/12$.

Poisson(λ): $p_X(k) = e^{-\lambda} \lambda^k / k!$, $k = 0, 1, \dots$ $E[X] = \lambda$, $\text{Var}(X) = \lambda$. Good approx for Binomial when n large, p small, $\lambda = np$.

8.3 Functions of Random Variables (Sec. 2.2)

If $Y = g(X)$: $p_Y(y) = \sum_{\{x : g(x) = y\}} p_X(x)$.

Example: X uniform on $\{0, \dots, 9\}$, $Y = X \bmod 3$: $p_Y(0) = P(X \in \{0, 3, 6, 9\}) = 4/10$; $p_Y(1) = 3/10$; $p_Y(2) = 3/10$.

$Z = 5 \bmod (X + 1)$: compute $5 \bmod k$ for each $k = 1, \dots, 10$ and aggregate.

X = product of heads and tails in n flips: $X = k(n - k)$ where $k \sim \text{Bin}(n, p)$. $P(X = 0) = p^n + (1-p)^n$.

8.4 Expected Value (Mean) (Sec. 2.3)

$E[X]=\sum_x x\,p_X(x)$ (weighted average of values).

LOTUS (Expected Value Rule): $E[g(X)]=\sum_x g(x)\,p_X(x)$.

Linearity (always holds): $E[aX+b]=aE[X]+b$. $E[X+Y]=E[X]+E[Y]$. $E[\sum_i X_i]=\sum_i E[X_i]$ (even if dependent).

$E[XY]=E[X]\,E[Y]$ **only when** X,Y independent.

n**th moment:** $E[X^n]=\sum_x x^n p_X(x)$.

8.5 Variance & Standard Deviation (Sec. 2.3)

$\mathrm{Var}(X)=E[(X-E[X])^2]=E[X^2]-(E[X])^2$.

$\mathrm{Var}(aX+b)=a^2\,\mathrm{Var}(X)$ (shift doesn't change variance).

$\sigma_X=\sqrt{\mathrm{Var}(X)}$. $\mathrm{Var}(X)\geq 0$; $=0$ iff X is constant.

Independent: $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$ (extends to n mutually indep. r.v.'s).

Variance of Binomial via indicator decomposition: $X=X_1+\cdots+X_n$, each $X_i\sim\mathrm{Bern}(p)$, independent. $\mathrm{Var}(X)=\sum\mathrm{Var}(X_i)=np(1-p)$.

Computing Var **from PMF of** $Z=(X-\mu)^2$**:** Find PMF of Z, then $\mathrm{Var}(X)=E[Z]=\sum z\,p_Z(z)$.

9. Problem-Solving Strategies & Common Patterns

Complement: $P(\text{at least one})=1-P(\text{none})$. E.g. $P(\geq 1\text{ defective in sample})=1-P(0\text{ defective})$.

Conditioning (Total Prob): Partition the scenario, compute each conditional, sum.

Sequential Bayes: Observe evidence one piece at a time; posterior from step k becomes prior for step $k+1$.

Tree diagrams: Draw branches for each stage; multiply along paths (chain rule); add across paths for total prob.

Geometric series: $\sum_{n=0}^{\infty}r^n=\frac{1}{1-r}$ for $|r|<1$. $\sum_{n=0}^{N-1}r^n=\frac{1-r^N}{1-r}$.

Binomial theorem: $(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}$.

Exponential limit: $\lim_{n\to\infty}(1+\alpha/n)^n=e^{\alpha}$.

Best-of-$(2m-1)$ **series:** First to m wins. Team wins iff it wins exactly m out of first k games, last game being a win: $P(\text{win series})=\sum_{k=m}^{2m-1}\binom{k-1}{m-1}p^m(1-p)^{k-m}$. If $p>1/2$: prefer longer series (more games $\Rightarrow$ better team wins more often).

Win-by-2 match: Games paired into rounds of 2. Round decisive with prob $p^2+(1-p)^2$. Given decisive, player 1 wins with $\frac{p^2}{p^2+(1-p)^2}$. Overall: $P=\frac{p^2}{p^2+(1-p)^2}$ (geometric series on tied rounds).

Repetition coding: Send bit 3 times; majority decode. $P(\text{correct})=p^3+3p^2(1-p)$; $P(\text{error})=3p(1-p)^2+(1-p)^3$ where $p=$ prob bit transmitted correctly.

Successful transmission (ALOHA): n transmitters, each sends w.p. p independently. $P(\text{success})=np(1-p)^{n-1}$. Max at $p=1/n$: $P_{\max}=(1-1/n)^{n-1}\to 1/e\approx 0.368$. Over T slots: $P(\geq 1\text{ success in }T\text{ slots})=1-(1-P_s)^T$ where $P_s=np(1-p)^{n-1}$. Min T for $P\geq\alpha$: $T\geq\lceil\ln(1-\alpha)/\ln(1-P_s)\rceil$.

Phone/hat matching: Each phone scratched w.p. p indep. $P(\text{first }m\text{ get scratched})=p^m$ (indep. of assignment). $P(\text{exactly }m\text{ get scratched})=\binom{n}{m}p^m(1-p)^{n-m}$.

Multiple-choice Bayes: Student knows answer w.p. α, guesses otherwise. Given correct: $P(\text{knew}\,|\,\text{correct})=\frac{\alpha}{\alpha+(1-\alpha)/m}$ where $m=\#$ choices. With partial elimination, weight each scenario.

Judge vs. jury: Single judge correct w.p. p. Three-person majority jury (two competent w.p. p, one flips coin): $P(\text{jury correct})=p^2+p(1-p)+\frac{1}{2}[p(1-p)+(1-p)^2]$. Compare to p.

Power plants (indep. failures): Plant k fails w.p. p_k. Any one suffices: $P(\text{blackout})=\prod_k p_k$. Need ≥ 2 running: $P(\text{blackout})=\prod_k p_k+\sum_j(1-p_j)\prod_{k\neq j}p_k$.

Useful Identities

$\sum_{m=1}^{n}m=\frac{n(n+1)}{2}$; $\sum_{m=1}^{n}m^2=\frac{n(n+1)(2n+1)}{6}$; $\sum_{k=1}^{n}\frac{1}{k}\approx\ln n+0.577$.

$n!=\sqrt{2\pi n}\,(n/e)^n$ (Stirling's approx).

10. Quick Reference: Independence & Disjointness

	If A,B disjoint	**If A,B independent**
$P(A\cap B)$	$=0$	$=P(A)P(B)$
$P(A\cup B)$	$=P(A)+P(B)$	$=P(A)+P(B)-P(A)P(B)$
$P(A\,\|\,B)$	$=0$	$=P(A)$
$P(B\,\|\,A)$	$=0$	$=P(B)$
Intuition	B happening *rules out* A	B happening says *nothing* about A
Can be both?	Only if $P(A)=0$ or $P(B)=0$	Only if $P(A)=0$ or $P(B)=0$

Summary: Disjoint = strong negative dependence. Independent = no dependence. Both with $P>0$ is **impossible**. If conditioning increases prob ($P(A|B)>P(A)$), events are positively associated (not indep., not disjoint).

11. Worked Examples & Patterns from HW/Discussion

Conditional Probability Pitfalls

Two coins, both heads: $P(\text{HH}\,|\,\text{1st is H})=p$. But $P(\text{HH}\,|\,\text{at least one H})=\frac{p^2}{2p-p^2}\leq p$. The second condition is weaker, so the conditional prob is smaller.

Non-uniform die: If face k has prob $\alpha(k+1)$, then $\sum_{k=0}^{n-1}\alpha(k+1)=1$ gives $\alpha=\frac{2}{n(n+1)}$. For 12-sided die: $\alpha=1/78$. $P(k)=(k+1)/78$ for $k=0,\ldots,11$.

Drawing without replacement: Box with 3 crayons. Draw one, return only if cyan, draw again. $P(\text{2nd draw}=l\,|\,\text{1st draw}=k)$ depends on whether k was returned. Chain rule gives joint prob.

Counting & Combinatorial Problems

Teams from n people: Assign n people to r teams of size k ($n=rk$): Total assignments $=\frac{n!}{(k!)^r}$ (if teams labeled); divide by $r!$ if teams unlabeled.

"Power team" prob: $n=12$ people, 4 teams of 3. Prob that 3 of 4 special people land on same team: $\frac{\binom{4}{3}\cdot 4\cdot\binom{8}{2}\cdot\frac{9!}{(3!)^3}}{\frac{12!}{(3!)^4}}$.

Volleyball rosters: Choose 6 from 15 (6 women, 9 men). Exactly 2 women, 4 men: $\binom{6}{2}\binom{9}{4}$. At least 2 women: $\binom{15}{6}-\binom{9}{6}-\binom{6}{1}\binom{9}{5}$.

Rooks problem: 8 rooks on 8×8 board, no two share row or column. Choose 8 rows from 8 ($\binom{8}{8}$), assign columns (8! permutations): $P=8!/\binom{64}{8}\cdot\frac{1}{8!}\cdot 8!=\frac{8!}{\binom{64}{8}}$.

Bayes' Rule Applications

Radar/detection: Prior: $P(\text{alien})=0.05$. $P(\text{blip}\,|\,\text{alien})=0.99$, $P(\text{blip}\,|\,\text{no alien})=0.1$. $P(\text{alien}\,|\,\text{blip})=\frac{0.99\times 0.05}{0.99\times 0.05+0.1\times 0.95}=\frac{0.0495}{0.1445}\approx 0.343$. $P(\text{miss})=P(\text{no blip}\,|\,\text{alien})P(\text{alien})=0.01\times 0.05=0.0005$. $P(\text{false alarm})=P(\text{blip}\,|\,\text{no alien})P(\text{no alien})=0.1\times 0.95=0.095$.

Biased die (sequential Bayes): Two dice: standard (p) and loaded. Roll 3, then 6, then 5. After each roll, update $P(\text{standard})$ using Bayes. After seeing 5: $P(\text{standard}\,|\,5)=1$ (loaded die can't produce 5).

Monty Hall/Prisoner variant: 3 prisoners, 2 released. Gollum asks which other prisoner is released. If Gollum is to be released (prob 2/3), guard's answer is determined. If not (prob 1/3), guard picks uniformly. By Bayes, $P(\text{Gollum released}\,|\,\text{guard says }X)=2/3$ regardless.

Multiple choice Bayes: Knows answer w.p. 1/2 (picks correct). W.p. 1/4 eliminates 1 wrong (picks from 3). Otherwise guesses from 4. $P(\text{knew}\,|\,\text{correct})=\frac{1/2}{1/2+\frac{1}{4}\cdot\frac{1}{3}+\frac{1}{4}\cdot\frac{1}{4}}=\frac{1/2}{1/2+1/12+1/16}=\frac{24}{37}$.

Independence Problems

Sum of dice and individual rolls: $A=\{\text{sum}=12\}$, $B=\{\text{at least one 6}\}$, $C=\{\text{at least one 2}\}$. $P(A)=1/36$, $P(B)=11/36$, $P(A\cap B)=1/36=P(A)$. So $A\perp\!\!\!\perp B$? Check: $P(A)P(B)=11/1296\neq 1/36$. **Not independent.** $P(C)=11/36$, $P(A\cap C)=0\neq P(A)P(C)$. Not independent (disjoint with $P>0$!).

Proving $P(A_1^c\cup\cdots\cup A_n^c)=1-\prod P(A_i)$**:** By De Morgan: $A_1^c\cup\cdots\cup A_n^c=(A_1\cap\cdots\cap A_n)^c$. $P=(1-P(A_1\cap\cdots\cap A_n))=1-\prod P(A_i)$ by independence.

Random Variable Computations

Finding PMF constant: If $p_X(x)=x^2/a$ for $x\in\{-3,\ldots,3\}$: $\sum x^2/a=1\Rightarrow(9+4+1+0+1+4+9)/a=1\Rightarrow a=28$. $E[X]=\sum x\cdot x^2/28=0$ (by symmetry). $E[X^2]=\sum x^2\cdot x^2/28=\sum x^4/28=(81+16+1+0+1+16+81)/28=196/28=7$. $\mathrm{Var}(X)=E[X^2]-(E[X])^2=7-0=7$.

Score with random test: 3 tests (easy $p=0.9$, med $p=0.7$, hard $p=0.5$), 3 questions each, chosen uniformly. $p_X(k)=\frac{1}{3}\left[\binom{3}{k}(0.9)^k(0.1)^{3-k}+\binom{3}{k}(0.7)^k(0.3)^{3-k}+\binom{3}{k}(0.5)^k(0.5)^{3-k}\right]$.

Product r.v.: $X=H\cdot T$ where $H+T=n$ flips. $X=k(n-k)$ for k heads. $P(X=0)=P(H=0)+P(H=n)=(1-p)^n+p^n$. For $n=4$: possible X values are $0,3,4$ (from $k=0,1,3,4\to 0$; $k=2\to 4$; $k=1,3\to 3$).

Geometric & Series Applications

Verification $\sum_{k=1}^{\infty}(1-p)^{k-1}p=1$**:** $=p\sum_{j=0}^{\infty}(1-p)^j=p\cdot\frac{1}{1-(1-p)}=1$. ✓

$E[X]$ **for geometric:** $\sum_{k=1}^{\infty}k(1-p)^{k-1}p=\frac{p}{(1-(1-p))^2}=\frac{1}{p}$. Useful: $\sum_{k=1}^{\infty}kx^{k-1}=\frac{1}{(1-x)^2}$ for $|x|<1$.

CDF of geometric: $P(X\leq k)=1-(1-p)^k$. $P(X>k)=(1-p)^k$.

Negative binomial as sum: $X=X_1+\cdots+X_r$ where $X_i\overset{\text{iid}}{\sim}\mathrm{Geom}(p)$. $E[X]=r/p$, $\mathrm{Var}(X)=r(1-p)/p^2$.

Best-of-N Series Details

Best-of-7: First to 4 wins. If Canadiens win each game w.p. p: $P(\text{Rangers win})=\sum_{k=4}^{7}\binom{k-1}{3}(1-p)^4p^{k-4}$ $=\sum_{j=0}^{3}\binom{j+3}{3}(1-p)^4p^j$ where $j=k-4$.

Equivalently: $P(\text{Rangers win series})=(1-p)^4\sum_{j=0}^{3}\binom{j+3}{3}p^j$.

Key insight: If $p>1/2$ (Leafs better), they prefer *longer* series—more games let skill dominate luck. Best-of-5 $>$ best-of-3 $>$ single game for the better team.

Probability Bounds & Ranges

Given $P(A)$ and $P(B)$ only (no info on overlap): $P(A\cup B)\in[\max(P(A),P(B)),\min(1,P(A)+P(B))]$. $P(A\cap B)\in[\max(0,P(A)+P(B)-1),\min(P(A),P(B))]$.

Ex: $P(D)=0.13$, $P(M)=0.37$. $P(D\cup M)\in[0.37,\,0.50]$; $P(D\cap M)\in[0,\,0.13]$.

Conditional Independence

A and B conditionally independent given C: $P(A\cap B\,|\,C)=P(A\,|\,C)P(B\,|\,C)$. Conditional independence $\not\Rightarrow$ unconditional independence (and vice versa).

Infinite Intersections

$\bigcap_{n=0}^{\infty}A_n$ where $A_n=\{m\in\mathbb{N}:m\geq n\}$: every natural number is eventually excluded, so $\bigcap A_n=\varnothing$.

Spinner / Continuous Finite Models

Spinner uniform on $[0,2\pi)$. Colors in quadrants: $P(\text{any color})=1/4$. Finite model: $\Omega=\{R,G,Y,B\}$, $P(\{c\})=1/4$. Infinite: $\Omega=[0,2\pi)$, $P=\theta/(2\pi)$. $P(\text{not yellow and not red})=P(\{G,B\})=1/2$.

Key Checks Before Answering

- Do probabilities sum to 1? ($\sum p_X(x)=1$.)
- Is the sample space correct and complete?
- Did you use the right formula (with vs. without replacement)?
- For independence: did you check $P(A\cap B)=P(A)P(B)$ (not just $P(A\,|\,B)$)?
- For Bayes: did you use total probability in the denominator?
- For counting: ordered or unordered? with or without replacement?