Course Notes Summary

Focus Area: Summary of main formulas please

Source:	BayesTheorem.pdf
Generated:	2025-09-05 16:25:15
Summary Type:	Comprehensive

Prepared for quick revision and reference

Use this sheet as a step-by-step guide when solving problems.

1 Key definitions

1.1 Conditional probability

- $P(B \mid A)$ = probability that B occurs given A has occurred.
- Formula: $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$, provided $P(A) > 0$.

1.2 Prior, likelihood, posterior, evidence

- Prior: $P(A)$ = initial probability of hypothesis A before new data.
- Likelihood: $P(B \mid A)$ = probability of observing data B assuming A is true.
- Evidence (marginal): $P(B)$ = total probability of observing B.
- Posterior: $P(A \mid B)$ = updated probability of hypothesis A after observing B.

1.3 Other useful facts

- Multiplication rule: $P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B)$.
- Independence: If A and B are independent, $P(A \cap B) = P(A)P(B)$ and $P(B \mid A) = P(B)$.

2 Main formulas

2.1 Basic conditional probability

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}.$$

2.2 Bayes' theorem (two events / hypothesis)

$$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)}.$$

2.3 Law of total probability (for partition $\{A_1, A_2, \ldots, A_n\}$)

$$P(B) = \sum_i P(A_i)\,P(B \mid A_i).$$

Use this as the denominator in Bayes' theorem when hypotheses form a partition.

2.4 Generalized Bayes' theorem (n hypotheses $A_1, \ldots, A_n$)

$$P(A_k \mid B) = \frac{P(A_k)\,P(B \mid A_k)}{\sum_i P(A_i)\,P(B \mid A_i)}.$$

2.5 Frequency/count form (intuitive/table method)

- Given counts, $P(A \mid B) = \dfrac{\#(A \cap B)}{\#(B)}$.
- To use probabilities, assume total N and compute cell counts: $\#(A \cap B) = N \times P(A) \times P(B \mid A)$.

2.6 Symmetry relation

$$P(A \mid B)\, P(B) = P(B \mid A)\, P(A).$$

3 Short worked examples

3.1 Cigar-smoking example (Orange County)

1. Given: $P(\text{Male}) = 0.51$, $P(\text{Female}) = 0.49$.
2. Given: $P(\text{Cigar} \mid \text{Male}) = 0.095$, $P(\text{Cigar} \mid \text{Female}) = 0.017$.
3. Compute numerator: $P(\text{Male})P(\text{Cigar} \mid \text{Male}) = 0.51 \times 0.095 = 0.04845$.
4. Compute denominator: $0.04845 + 0.49 \times 0.017 = 0.04845 + 0.00833 = 0.05678$.
5. Posterior: $P(\text{Male} \mid \text{Cigar}) = \dfrac{0.04845}{0.05678} \approx 0.853$.

3.2 ELT (manufacturers) defective example

1. Given priors: $P(A) = 0.80$, $P(B) = 0.15$, $P(C) = 0.05$.
2. Given defect rates: $P(D \mid A) = 0.04$, $P(D \mid B) = 0.06$, $P(D \mid C) = 0.09$.
3. Numerator for A: $0.80 \times 0.04 = 0.032$.
4. Denominator: $0.032 + 0.15 \times 0.06 + 0.05 \times 0.09 = 0.032 + 0.009 + 0.0045 = 0.0455$.
5. Posterior: $P(A \mid D) = \dfrac{0.032}{0.0455} \approx 0.703$.

3.3 Table method (same cigar example with $N = 100{,}000$)

1. Compute counts: males $= 51{,}000$, cigar-smoking males $= 0.095 \times 51{,}000 = 4{,}845$.
2. females $= 49{,}000$, cigar-smoking females $= 0.017 \times 49{,}000 = 833$.
3. total cigar-smokers $= 4{,}845 + 833 = 5{,}678$.
4. $P(\text{Male} \mid \text{Cigar}) = 4{,}845/5{,}678 \approx 0.853$.

4 Important relationships & connections

- Posterior $\propto$ Prior $\times$ Likelihood: $P(A \mid B) \propto P(A)P(B \mid A)$. The denominator normalizes over all hypotheses.
- Law of total probability supplies the evidence term required for normalization.
- Bayes' theorem updates beliefs when new data arrives. Use priors that partition the sample space.
- Frequency approach (assume an N) reduces algebra errors and aids intuition.
- Independence simplifies many computations, but do not assume independence unless justified.

5 Quick checklist for applying Bayes' theorem

1. Identify hypotheses that partition the sample space: $\{A_1, \ldots, A_n\}$.
2. Obtain priors $P(A_i)$ for each hypothesis.
3. Obtain likelihoods $P(B \mid A_i)$ for observed data B.
4. Compute evidence $P(B) = \sum_i P(A_i)P(B \mid A_i)$.
5. Compute posterior for target hypothesis: $P(A_k \mid B) = \dfrac{P(A_k)P(B \mid A_k)}{P(B)}$.
6. Optionally verify with a contingency table using an assumed total N.

6 Common pitfalls

- Forgetting to normalize (omitting $P(B)$).
- Using a non-partitioned or incomplete set of hypotheses in the denominator.
- Confusing $P(B \mid A)$ (likelihood) with $P(A \mid B)$ (posterior).
- Applying independence incorrectly.

7 One-line summaries

- Bayes (two events): $P(A \mid B) = \dfrac{P(B \mid A)P(A)}{P(B)}$.
- Generalized Bayes: $P(A_k \mid B) = \dfrac{P(A_k)P(B \mid A_k)}{\sum_i P(A_i)P(B \mid A_i)}$.
- Frequency method: $P(A \mid B) = \dfrac{\#(A \cap B)}{\#(B)}$.