Course Notes Summary

Focus Area: Summary of main formulas please

Source:	BayesTheorem.pdf
Generated:	2025-09-05 15:53:31
Summary Type:	Comprehensive

Prepared for quick revision and reference

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

1 Section 1.

1.1 Subsection 1.1

- Key definitions
  - Prior probability: initial probability before new information (e.g., $P(A)$).
  - Posterior probability: revised probability after incorporating new information (e.g., $P(A \mid B)$).
  - Likelihood: probability of observed evidence given a hypothesis (e.g., $P(B \mid A)$).
  - Marginal (or total) probability: overall probability of evidence (e.g., $P(B)$).
  - Conditional probability: probability of one event given another ($P(B \mid A)$).
- Fundamental formulas
  - Conditional probability (definition):
    * $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$
  - Product rule (relates joint and conditional probabilities):
    * $P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B)$
  - Independence criterion:
    * A and B independent $\Leftrightarrow$ $P(A \cap B) = P(A)P(B)$
    * Equivalently $P(B \mid A) = P(B)$ (when independent)

1.2 Subsection 1.2

- Bayes' theorem (two events)
  - $P(A \mid B) = \dfrac{P(B \mid A)P(A)}{P(B)}$
  - Use when you want posterior $P(A \mid B)$ and you know prior $P(A)$ and likelihood $P(B \mid A)$.
- Law of total probability (two-event form):
  - $P(B) = P(A)P(B \mid A) + P(A^c)P(B \mid A^c)$
- Generalized Bayes (partition $A_1, \ldots, A_n$)
  - Law of total probability (general):
    * $P(B) = \sum_i P(A_i)P(B \mid A_i)$
  - Bayes' theorem (generalized):
    * $P(A_k \mid B) = \dfrac{P(A_k)P(B \mid A_k)}{\sum_i P(A_i)P(B \mid A_i)}$
  - Compact relationship (proportional form):
    * $P(A_k \mid B) \propto P(A_k)\, P(B \mid A_k)$ (normalize by sum over k)

2 Section 2.

2.1 Subsection 2.1

- Stepwise procedure to apply Bayes' theorem (ordered)
  1. List hypotheses (partition) $A_1, \ldots, A_n$ and their priors $P(A_i)$.
  2. For each hypothesis, list the likelihood $P(B \mid A_i)$.
  3. Compute marginal $P(B) = \sum_i P(A_i)P(B \mid A_i)$.
  4. Compute posterior $P(A_k \mid B) = \dfrac{P(A_k)P(B \mid A_k)}{P(B)}$.
- Intuitive (frequency/table) method
  - Assume a convenient total N (e.g., 1000 or 100000).
  - Convert priors to counts: $\text{count}(A_i) = N \cdot P(A_i)$.
  - For each A_i, compute $\text{count}(A_i \wedge B) = \text{count}(A_i) \cdot P(B \mid A_i)$.
  - Marginal $\text{count}(B) = \sum_i \text{count}(A_i \wedge B)$.
  - Posterior $= \dfrac{\text{count}(A_k \wedge B)}{\text{count}(B)}$.
- Short example 1 (Orange County cigar smokers)
  - Given: $P(\text{male}) = 0.51$, $P(\text{cigar} \mid \text{male}) = 0.095$, $P(\text{cigar} \mid \text{female}) = 0.017$.
  - Compute $P(\text{male} \mid \text{cigar})$:
    * Numerator $= 0.095 \times 0.51 = 0.04845$
    * Denominator $= 0.04845 + 0.017 \times 0.49 = 0.04845 + 0.00833 = 0.05678$
    * $P(\text{male} \mid \text{cigar}) = \dfrac{0.04845}{0.05678} \approx 0.853$
- Short example 2 (ELT manufacturers and defectives)
  - Given: $P(A) = 0.80$, $P(B) = 0.15$, $P(C) = 0.05$; defect rates $P(D \mid A) = 0.04$, $P(D \mid B) = 0.06$, $P(D \mid C) = 0.09$.
  - Compute $P(A \mid D)$:
    * Numerator $= 0.80 \times 0.04 = 0.032$
    * Denominator $= 0.032 + 0.15 \times 0.06 + 0.05 \times 0.09 = 0.032 + 0.009 + 0.0045 = 0.0455$
    * $P(A \mid D) = \dfrac{0.032}{0.0455} \approx 0.703$
  - Equivalent frequency-table example ($N = 10{,}000$): defective counts $A = 320$, $B = 90$, $C = 45$, total defective $= 455 \Rightarrow P(A \mid D) = 320/455 \approx 0.703$.

2.2 Subsection 2.2

- Important relationships and connections
  - Prior $\to$ Likelihood $\to$ Posterior: Bayes updates prior by weighting it with the likelihood of observed evidence.
  - Denominator role: $P(B)$ is the marginal likelihood; it ensures posterior probabilities sum to 1 across hypotheses.
  - Proportionality: Posterior is proportional to prior $\times$ likelihood; normalization uses total probability.

  - Generalization: Bayes works with any finite partition of hypotheses; use law of total probability for the denominator.

- Practical tips and common pitfalls
  - Tip: Use the frequency/table method when formulas feel error-prone; it is less error-prone and intuitive.
  - Tip: Keep track of which event is the hypothesis (A_i) and which is the evidence (B).
  - Pitfall: Confusing $P(B \mid A)$ (likelihood) with $P(A \mid B)$ (posterior).
  - Pitfall: Forgetting to include all partition events in the denominator (use full partition).

- Compact summary of main formulas (for quick revision)
  - $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$
  - $P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B)$
  - $P(A \mid B) = \dfrac{P(B \mid A)P(A)}{P(B)}$
  - $P(B) = \sum_i P(A_i)P(B \mid A_i)$
  - $P(A_k \mid B) = \dfrac{P(A_k)P(B \mid A_k)}{\sum_i P(A_i)P(B \mid A_i)}$