AP Statistics Unit 4 – Quick Notes

1. Probability Basics

Event: A set of outcomes from a random process. **Sample space** (S): All possible outcomes. **Notation:** $P(A)$, A^c (complement), $A \cap B$ (both occur), $A \cup B$ (at least one occurs).

Rules:

- $0 \leq P(A) \leq 1$ (probability is always between 0 and 1)
- $P(S) = 1$ (the probability of the whole sample space is 1)
- **Addition Rule:**

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

  If mutually exclusive: $P(A \cap B) = 0$, so $P(A \cup B) = P(A) + P(B)$.

Example: Rolling a die: $P(\text{even or prime}) = P(\text{even}) + P(\text{prime}) - P(\text{even and prime})$.

2. Conditional Probability

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

"Given" means the sample space is restricted to cases where B has occurred.

Independence: A and B are independent if $P(A|B) = P(A)$ or equivalently $P(A \cap B) = P(A)P(B)$.

Example: If $P(\text{hockey}|\text{Canada}) = 0.67$ but $P(\text{hockey}) = 0.67$, they are independent.

3. Law of Total Probability & Bayes

Law of Total Probability:

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

Breaks A into cases based on whether B happens.

Bayes' Theorem:

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

Reverses conditional probabilities; useful for diagnostic testing.

Example: Given test accuracy and disease rate, find $P(\text{disease}|\text{positive})$.

4. Counting

Multiplication Rule: If first step has m outcomes and second has n, total $= m \times n$.

Permutations (order matters):

$$P(n, r) = \frac{n!}{(n-r)!}$$

Combinations (order doesn't matter):

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Example: Choosing 3 students from 10 $= \binom{10}{3}$.

5. Simulation Steps

1. State problem clearly.
2. Identify assumptions (e.g., independence, fixed p).
3. Assign numbers to outcomes.
4. Simulate many trials (random digits, calculator, computer).
5. Estimate probability from relative frequency.

Simulations approximate probabilities when theory is complex.

6. Random Variables

A **random variable** assigns a number to each outcome.

Expected Value (mean):

$$E[X] = \sum x_i P(x_i)$$

Variance:

$$\text{Var}(X) = E[(X - \mu_x)^2] = \sum (x_i - \mu)^2 P(x_i) = E[X^2] - E[X]^2$$

Std. deviation: $\sigma = \sqrt{\text{Var}(X)}$.

Example: Payoff with probabilities: multiply each outcome by its probability and sum.

7. Special Discrete Distributions

Binomial: Fixed n, success/failure, independent, constant p.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$\mu = np, \quad \sigma^2 = np(1-p)$.

Geometric: Trials until first success.

$$P(Y = k) = (1-p)^{k-1}p, \quad E[Y] = \frac{1}{p}$$

8. Continuous Distributions

Normal: $N(\mu, \sigma)$, use $z = \frac{x-\mu}{\sigma}$ and `normalcdf`.

Mean of sums/differences: $E(X \pm Y) = E(X) \pm E(Y)$. Variance (independent): $\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$.

Example: Two independent sample means: variances add.

9. Independence vs. Mutually Exclusive

- Mutually exclusive: $P(A \cap B) = 0$ (cannot occur together).
- Independent: $P(A \cap B) = P(A)P(B)$.
- Cannot be both if $P(A), P(B) > 0$.