Probability and sampling

AP STATISTICS DUMP SHEET

Achievable

Probability and random variables:

- **Basic probability rules:**
  $P(A^c) = 1 - P(A)$
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

- **Rules for independence:**
  $P(A \mid B) = P(A)$
  $P(A \cap B) = P(A)\,P(B)$

- **Mean and variance of discrete random variable:**
  $\mu = \sum x_i p_i$
  $\sigma^2 = \sum (x_i - \mu)^2 p_i$

- **Binomial distribution:**
  $P(X{=}k) = \binom{n}{k} p^k (1-p)^{n-k}$
  $\mu = np, \quad \sigma = \sqrt{np(1-p)}$

  binary, independent, consistent probability of success.

- **Geometric distribution:**
  $P(X{=}k) = (1-p)^{k-1} p$
  $\mu = \frac{1}{p}, \quad \sigma = \frac{\sqrt{1-p}}{p}$

- **Bayes' theorem:**
  $P(A \mid B) = \frac{P(\,B \mid A\,)\,P(A)}{P(B)}$

- **Variance of a random variable:**
  $\mathrm{Var}(X) = \sum x_i^{\,2} p_i - \mu^2$

- **Expected value rules:**
  $E(aX + b) = aE(X) + b$
  $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$

Sampling distributions:

- **Central limit theorem:**
  For large n (typically $n \geq 30$), the sampling distribution of $\bar{x}$ (the sample mean) is approximately normal regardless of the population shape if the observations are independent.

- **Unbiased estimator:** mean of the sampling distribution equals the parameter.

- **10% condition:** For sampling without replacement, the sample size must be less than or equal to 10% of the population to treat observations as independent in formulas.

- A statistic varies from sample to sample. This variation forms a sampling distribution with $\hat{p}$ as $\bar{x}$ unbiased estimators of p and μ.

- **Sampling distribution of a proportion:** $\hat{p} \sim N = \left(p, \sqrt{\frac{p(1-p)}{n}}\right)$

- **Conditions:** $np \geq 10, \quad n\,(1-p) \geq 10$
  and a random sample that is no more than 10% of the population.

- **Sampling distribution of a mean:** $\bar{x} \sim N = \left(\mu, \frac{\sigma}{\sqrt{n}}\right)$