Full Name ______________________ ID ______________________

Cheat Sheet for Probability and Statistics 00340058

1 Basic Probability

Sample Space and Events:
Sample space Ω: set of all possible outcomes
Event A: subset of sample space
Complement: $A^c = \{x \in \Omega : x \notin A\}$
Union: $A \cup B = \{x : x \in A \text{ or } x \in B\}$
Intersection: $A \cap B = \{x : x \in A \text{ and } x \in B\}$
Disjoint events: $A \cap B = \emptyset$

Probability Axioms:
For any event A:

$$0 \leq P(A) \leq 1$$
$$P(\Omega) = 1$$
$$P(\emptyset) = 0$$
$$P(A^c) = 1 - P(A)$$

Additive Rules:
General additive rule:

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

For mutually exclusive events ($A \cap B = \emptyset$):

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

Conditional Probability:

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

Multiplication rule:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Independence: A and B are independent if

$$P(A \cap B) = P(A)P(B)$$

or equivalently $P(A|B) = P(A)$

Law of Total Probability:
If $B_1, B_2, \ldots, B_k$ form a partition of Ω:

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

Bayes' Theorem:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{k} P(A|B_j)P(B_j)}$$

Counting Principles:
Multiplication rule: $n_1 \times n_2 \times \cdots \times n_k$ ways
Permutations: $P(n, r) = \frac{n!}{(n-r)!}$
Combinations: $\binom{n}{r} = \frac{n!}{r!(n-r)!}$
Circular permutations: $(n-1)!$
With repetition: $\frac{n!}{n_1! n_2! \cdots n_k!}$

2 Random Variables

Discrete Random Variables:
Probability mass function (PMF): $f(x) = P(X = x)$
Properties:

$$f(x) \geq 0 \text{ for all } x$$
$$\sum_{\text{all } x} f(x) = 1$$

Cumulative distribution function (CDF):

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

Continuous Random Variables:
Probability density function (PDF): $f(x)$
Properties:

$$f(x) \geq 0 \text{ for all } x$$
$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Cumulative distribution function:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

$$P(a < X < b) = \int_{a}^{b} f(x)dx = F(b) - F(a)$$

Expected Value (Mean):
Discrete: $E(X) = \mu = \sum_{\text{all } x} x \cdot f(x)$
Continuous: $E(X) = \mu = \int_{-\infty}^{\infty} x \cdot f(x)dx$

Properties of expectation:

$$E(c) = c$$
$$E(cX) = cE(X)$$
$$E(X + Y) = E(X) + E(Y)$$
$$E(aX + b) = aE(X) + b$$

Variance and Standard Deviation:

$$\text{Var}(X) = \sigma^2 = E[(X-\mu)^2] = E(X^2) - [E(X)]^2$$

Computing $E(X^2)$:

- Discrete: $E(X^2) = \sum_{\text{all } x} x^2 \cdot f(x)$
- Continuous: $E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx$

Alternative formulas:

- Discrete: $\sigma^2 = \sum_{\text{all } x}(x-\mu)^2 f(x)$
- Continuous: $\sigma^2 = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$

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

Properties of variance:

$$\begin{aligned} &\text{Var}(c) = 0 \\ &\text{Var}(cX) = c^2\text{Var}(X) \\ &\text{Var}(aX+b) = a^2\text{Var}(X) \end{aligned}$$

Variance Calculation Example:
Discrete distribution: X takes values 1, 2, 3 with probabilities 0.2, 0.5, 0.3

Step 1: $E(X) = 1(0.2) + 2(0.5) + 3(0.3) = 2.1$
Step 2: $E(X^2) = 1^2(0.2) + 2^2(0.5) + 3^2(0.3) = 0.2 + 2.0 + 2.7 = 4.9$
Step 3: $\text{Var}(X) = E(X^2) - [E(X)]^2 = 4.9 - (2.1)^2 = 4.9 - 4.41 = 0.49$

Covariance and Correlation:
Covariance: $\text{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$
Alternative form: $\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$

Correlation coefficient:

$$\rho = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

Properties: $-1 \leq \rho \leq 1$
If X and Y are independent, then $\text{Cov}(X,Y) = 0$

3 Discrete Distributions

Binomial Distribution:
$X \sim \text{Binomial}(n,p)$
Parameters:

- n: number of independent trials (fixed)
- p: probability of success on each trial (constant)
- x: number of successes observed ($0 \leq x \leq n$)

PMF: $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \ldots, n$
Mean: $\mu = np$
Variance: $\sigma^2 = np(1-p)$

Use when: Fixed number of independent trials, each with probability p of success

Hypergeometric Distribution:
$X \sim \text{Hypergeometric}(N, K, n)$
Parameters:

- N: total population size
- K: number of success states in population
- n: number of draws (sample size)
- x: number of observed successes in sample

PMF: $f(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$
where $\max(0, n-(N-K)) \leq x \leq \min(n, K)$

Mean: $\mu = n\frac{K}{N}$
Variance: $\sigma^2 = n\frac{K}{N}\left(1 - \frac{K}{N}\right)\frac{N-n}{N-1}$

Use when: Sampling without replacement from finite population

Geometric Distribution:
$X \sim \text{Geometric}(p)$ (number of trials until first success)
Parameters:

- p: probability of success on each trial (constant)
- x: trial number on which first success occurs ($x = 1, 2, 3, \ldots$)

PMF: $f(x) = p(1-p)^{x-1}$, $x = 1, 2, 3, \ldots$
Mean: $\mu = \frac{1}{p}$
Variance: $\sigma^2 = \frac{1-p}{p^2}$

Memoryless property: $P(X > s+t | X > s) = P(X > t)$

Negative Binomial Distribution:
$X \sim \text{NegBinomial}(r, p)$ (number of trials until r-th success)
Parameters:

- r: target number of successes (positive integer)
- p: probability of success on each trial (constant)
- x: trial number on which r-th success occurs ($x = r, r+1, \ldots$)

PMF: $f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$, $x = r, r+1, \ldots$
Mean: $\mu = \frac{r}{p}$
Variance: $\sigma^2 = \frac{r(1-p)}{p^2}$

Special case: Geometric is NegBinomial with $r = 1$

Poisson Distribution:
$X \sim \text{Poisson}(\lambda)$
Parameters:

- λ: average rate of events per unit time/space ($\lambda > 0$)
- x: number of events observed in the unit ($x = 0, 1, 2, \ldots$)

PMF: $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$, $x = 0, 1, 2, \ldots$
Mean: $\mu = \lambda$
Variance: $\sigma^2 = \lambda$

Use when: Counting rare events in time/space
Approximation to binomial when n is large, p is small, $np = \lambda$

Binomial Example:
Component survival probability = 0.75. Find P(exactly 2 of 4 survive):

$$P(X = 2) = \binom{4}{2}(0.75)^2(0.25)^2 = 6 \times \frac{9}{256} = \frac{27}{128} \approx 0.211$$

Poisson Example:
Average 3 accidents per month at intersection.
Find P(at most 2 accidents next month):

$$P(X \leq 2) = e^{-3}\left(\frac{3^0}{0!} + \frac{3^1}{1!} + \frac{3^2}{2!}\right)$$

$$= e^{-3}(1 + 3 + 4.5) = 8.5e^{-3} \approx 0.423$$

Hypergeometric Example:
Batch of 20 components: 5 defective, 15 good. Sample 4 without replacement.
Find P(exactly 1 defective):

$$P(X = 1) = \frac{\binom{5}{1}\binom{15}{3}}{\binom{20}{4}} = \frac{5 \times 455}{4845} = \frac{2275}{4845} \approx 0.469$$

4 Continuous Distributions

Uniform Distribution:
$X \sim \text{Uniform}(a, b)$
Parameters:

- a: lower bound of the interval
- b: upper bound of the interval ($b > a$)
- x: observed value ($a \leq x \leq b$)

PDF: $f(x) = \frac{1}{b-a}$, $a \leq x \leq b$
CDF: $F(x) = \frac{x-a}{b-a}$, $a \leq x \leq b$
Mean: $\mu = \frac{a+b}{2}$
Variance: $\sigma^2 = \frac{(b-a)^2}{12}$

Exponential Distribution:
$X \sim \text{Exponential}(\lambda)$
Parameters:

- λ: rate parameter ($\lambda > 0$)
- x: observed time or distance ($x \geq 0$)

PDF: $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$
CDF: $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$
Mean: $\mu = \frac{1}{\lambda}$
Variance: $\sigma^2 = \frac{1}{\lambda^2}$

Memoryless property: $P(X > s + t | X > s) = P(X > t)$
Use for: Waiting times, reliability analysis

Normal Distribution:
$X \sim N(\mu, \sigma^2)$
Parameters:

- μ: mean (location parameter)
- σ^2: variance; σ is standard deviation (scale parameter)
- x: observed value ($-\infty < x < \infty$)

PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Standard normal: $Z \sim N(0, 1)$
Standardization: $Z = \frac{X-\mu}{\sigma}$

Properties:

$$P(-1.96 < Z < 1.96) = 0.95$$
$$P(-2.58 < Z < 2.58) = 0.99$$
$$P(-1.64 < Z < 1.64) = 0.90$$

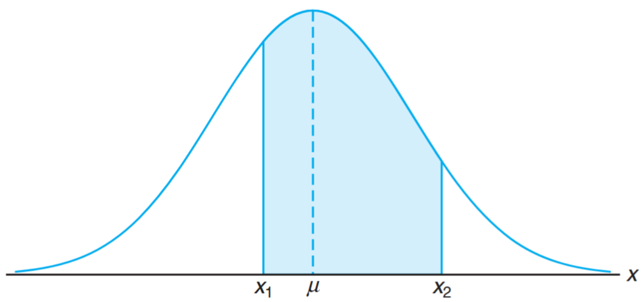

Figure 1: Standard normal distribution with common areas

Gamma Distribution:
$X \sim \text{Gamma}(\alpha, \beta)$
Parameters:

- α: shape parameter ($\alpha > 0$)
- β: rate parameter ($\beta > 0$)
- x: observed value ($x > 0$)

PDF: $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$, $x > 0$
Mean: $\mu = \frac{\alpha}{\beta}$
Variance: $\sigma^2 = \frac{\alpha}{\beta^2}$

Special cases:

- Exponential: $\alpha = 1$
- Chi-square: $\alpha = \nu/2$, $\beta = 1/2$

Normal Distribution Example:
IQ scores: $\mu = 100$, $\sigma = 15$. Find P(IQ between 85 and 115):

$$P(85 < X < 115) = P\left(\frac{85 - 100}{15} < Z < \frac{115 - 100}{15}\right)$$

$$= P(-1 < Z < 1) = 0.6826$$

Exponential Example:
Component lifetime: exponential with $\lambda = 0.02$ per hour.
Find P(lasts more than 50 hours):

$$P(X > 50) = 1 - F(50) = 1 - (1 - e^{-0.02 \times 50})$$
$$= e^{-1} \approx 0.368$$

Mean lifetime: $\mu = \frac{1}{0.02} = 50$ hours

5 Sampling Distributions

Sample Mean Distribution:
If $X_1, X_2, \ldots, X_n$ are iid from population with mean μ and variance σ^2:
Sample mean: $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Properties:

$$E(\overline{X}) = \mu$$
$$\text{Var}(\overline{X}) = \frac{\sigma^2}{n}$$
$$\text{SE}(\overline{X}) = \frac{\sigma}{\sqrt{n}}$$

Central Limit Theorem:
For large n (typically $n \geq 30$):

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0, 1)$$

Or equivalently: $\overline{X} \xrightarrow{d} N\left(\mu, \frac{\sigma^2}{n}\right)$

Sum of sample: $\sum_{i=1}^{n} X_i \xrightarrow{d} N(n\mu, n\sigma^2)$

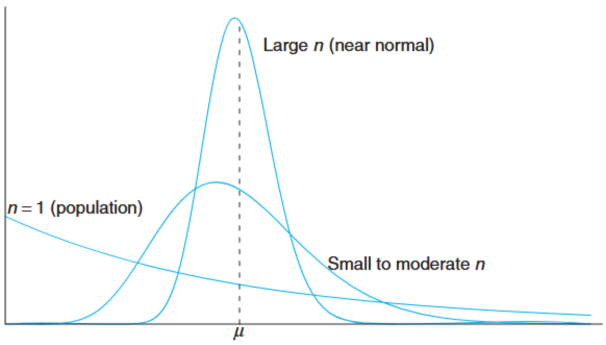

Figure 2: Central Limit Theorem illustration showing how sample mean distribution approaches normality

Chi-Square Distribution:
$\chi^2 \sim \chi^2_\nu$ with ν degrees of freedom

If $X \sim N(\mu, \sigma^2)$ and S^2 is sample variance:

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Mean: $E(\chi^2_\nu) = \nu$
Variance: $\text{Var}(\chi^2_\nu) = 2\nu$

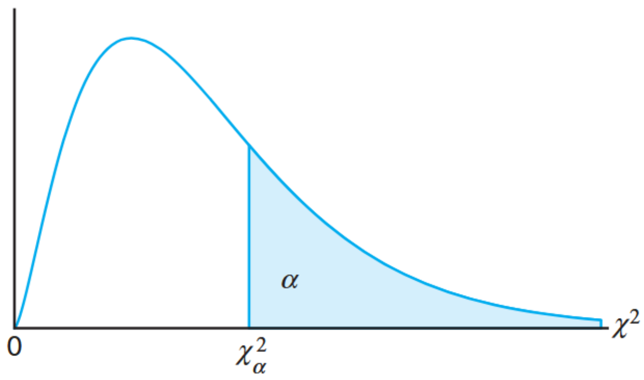

Figure 3: Chi-square distribution for different degrees of freedom

t-Distribution:
$t \sim t_\nu$ with ν degrees of freedom
Parameters:

- ν: degrees of freedom (positive integer)
- t: observed t-statistic ($-\infty < t < \infty$)

If $\overline{X}$ and S are sample mean and standard deviation from $N(\mu, \sigma^2)$:

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

As $\nu \to \infty$, $t_\nu \to N(0, 1)$
Symmetric around 0, heavier tails than normal

Central Limit Theorem Example:
Population: $\mu = 25$, $\sigma = 4$. Sample size $n = 36$.
Find P(sample mean exceeds 26):

$$P(\overline{X} > 26) = P\left(Z > \frac{26 - 25}{4/\sqrt{36}}\right)$$
$$= P(Z > 1.5) = 0.067$$

F-Distribution:
$F \sim F_{\nu_1, \nu_2}$ with ν_1 and ν_2 degrees of freedom
Parameters:

- ν_1: numerator degrees of freedom (positive integer)
- ν_2: denominator degrees of freedom (positive integer)
- F: observed F-statistic ($F \geq 0$)

If S_1^2 and S_2^2 are sample variances from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$:

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n_1-1, n_2-1}$$

Always positive, right-skewed

6 Confidence Intervals

Mean (σ known):
$100(1-\alpha)\%$ CI for μ:

$$\overline{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Sample size for margin of error E:

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$$

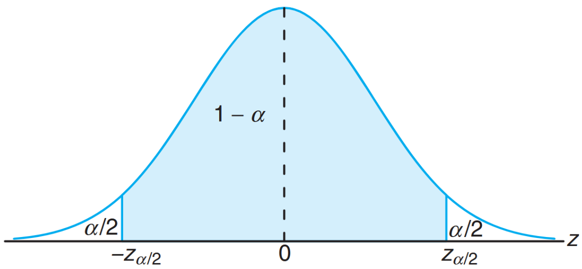

Figure 4: Confidence interval interpretation

Mean (σ unknown):
$100(1-\alpha)\%$ CI for μ (small sample, normal population):

$$\overline{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

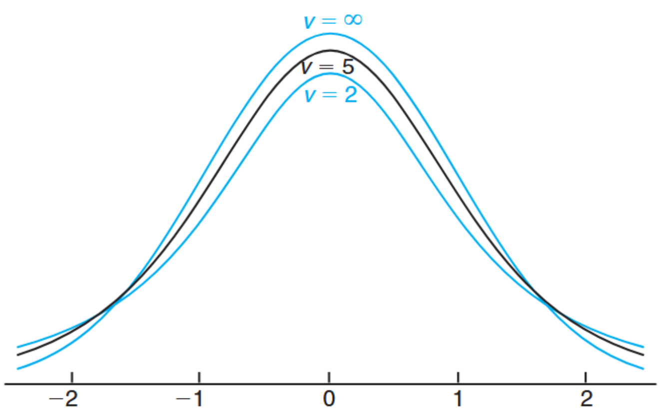

Figure 5: t-distribution for different degrees of freedom

Proportion:
$100(1-\alpha)\%$ CI for p (large sample):

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

where $\hat{p} = \frac{x}{n}$
Sample size: $n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p)$
Conservative: use $p = 0.5$

Variance:
$100(1-\alpha)\%$ CI for σ^2 (normal population):

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}$$

Difference of Means:
Independent samples, σ_1, σ_2 known:

$$(\overline{x}_1 - \overline{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$\sigma_1 = \sigma_2 = \sigma$ unknown, equal variances:

$$(\overline{x}_1 - \overline{x}_2) \pm t_{\alpha/2,n_1+n_2-2}\, s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where $s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

7 Hypothesis Testing

Test Components:
Null hypothesis: H_0 (assumed true)
Alternative hypothesis: H_1 or H_a
Test statistic: calculated from sample data
P-value: probability of observed result under H_0
Significance level: α (Type I error rate)

Decision rules:

- Reject H_0 if p-value $< \alpha$
- Reject H_0 if test statistic in critical region

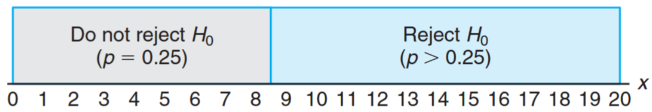

Figure 6: Critical region illustration for hypothesis testing

Errors in Testing:
Type I Error: Reject true H_0, $P(\text{Type I}) = \alpha$
Type II Error: Fail to reject false H_0, $P(\text{Type II}) = \beta$
Power: $1 - \beta = P(\text{reject false } H_0)$

Trade-off: Decreasing α increases β
Increase sample size to decrease both errors

Hypothesis Test Example:
Testing $H_0 : \mu = 68$ kg vs $H_1 : \mu \neq 68$ kg
Sample: $n = 36$, $\overline{x} = 67.5$, $s = 3.6$

Test statistic: $t = \frac{67.5-68}{3.6/\sqrt{36}} = \frac{-0.5}{0.6} = -0.833$

Critical values: $\pm t_{0.025,35} = \pm 2.03$

Since $|t| = 0.833 < 2.03$, fail to reject H_0

Type II Error Example:
Testing $H_0 : \mu = 50$ vs $H_1 : \mu = 52$ with $\sigma = 5$, $n = 25$, $\alpha = 0.05$

Critical value: $\overline{x}_c = 50 + 1.64 \times \frac{5}{\sqrt{25}} = 51.64$

When true mean is 52:

$$\beta = P(\overline{X} < 51.64 | \mu = 52) = P\left(Z < \frac{51.64 - 52}{1}\right)$$

$$= P(Z < -0.36) = 0.359$$

Power $= 1 - \beta = 0.641$

8 Tests for Mean (sigma known)

Tests for Mean (sigma known):
Test statistic: $Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$

Critical values:

- Two-tailed ($H_1 : \mu \neq \mu_0$): $\pm z_{\alpha/2}$
- Upper-tailed ($H_1 : \mu > \mu_0$): z_α
- Lower-tailed ($H_1 : \mu < \mu_0$): $-z_\alpha$

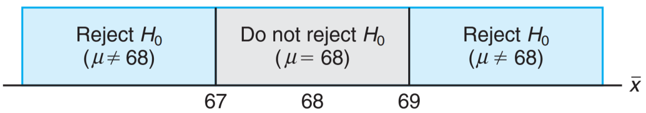

Figure 7: Critical regions for different alternative hypotheses

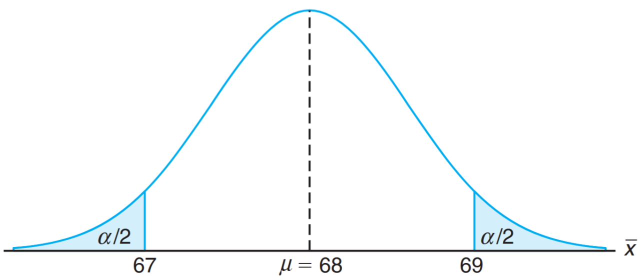

Figure 8: Two-tailed test critical region

Tests for Mean (σ unknown):
Test statistic: $T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$

Critical values: Replace z with $t_{\alpha,n-1}$ in previous formula

Tests for Proportion:
Test statistic: $Z = \frac{\hat{P} - p_0}{\sqrt{p_0(1-p_0)/n}}$
where $\hat{p} = \frac{x}{n}$, large sample required

Tests for Variance:
Test statistic: $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$

Critical values depend on alternative:

- $H_1 : \sigma^2 \neq \sigma_0^2$: $\chi^2_{\alpha/2,n-1}$ and $\chi^2_{1-\alpha/2,n-1}$
- $H_1 : \sigma^2 > \sigma_0^2$: $\chi^2_{\alpha,n-1}$
- $H_1 : \sigma^2 < \sigma_0^2$: $\chi^2_{1-\alpha,n-1}$

Two-Sample Tests:
Difference of means (equal variances):

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

Equality of variances:

$$F = \frac{S_1^2}{S_2^2} \sim F_{n_1-1,n_2-1}$$

Difference of proportions:

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$$

where $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$

9 Linear Regression

Simple Linear Regression Model:

$$Y = \beta_0 + \beta_1 x + \epsilon$$

where $\epsilon \sim N(0, \sigma^2)$

Least squares estimates:

$$\hat{\beta}_1 = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} = \frac{S_{xy}}{S_{xx}}$$

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}$$

Fitted line: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

Regression Calculations:
Key computational formulas:

$$S_{xx} = \sum (x_i - \overline{x})^2 = \sum x_i^2 - n\overline{x}^2$$
$$S_{yy} = \sum (y_i - \overline{y})^2 = \sum y_i^2 - n\overline{y}^2$$
$$S_{xy} = \sum (x_i - \overline{x})(y_i - \overline{y}) = \sum x_i y_i - n\overline{xy}$$

Where:

- S_{xx}: sum of squares of deviations in x
- S_{yy}: sum of squares of deviations in y
- S_{xy}: sum of cross products of deviations
- $\overline{x} = \frac{1}{n}\sum x_i$, $\overline{y} = \frac{1}{n}\sum y_i$

Example: Data points: $(1,2), (2,3), (3,5)$

$$n = 3, \quad \overline{x} = \frac{1+2+3}{3} = 2, \quad \overline{y} = \frac{2+3+5}{3} = \frac{10}{3}$$
$$S_{xx} = (1-2)^2 + (2-2)^2 + (3-2)^2 = 1 + 0 + 1 = 2$$
$$S_{xy} = (1-2)(2-\frac{10}{3}) + (2-2)(3-\frac{10}{3}) + (3-2)(5-\frac{10}{3})$$
$$= (-1)(-\frac{4}{3}) + (0)(-\frac{1}{3}) + (1)(\frac{5}{3}) = \frac{4}{3} + \frac{5}{3} = 3$$
$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{3}{2} = 1.5$$

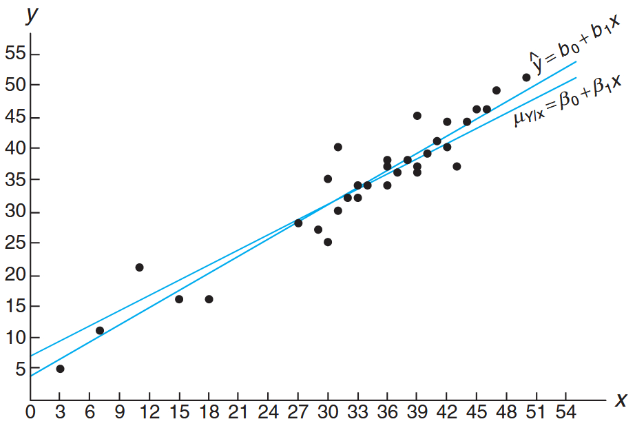

Figure 9: Scatter plot with fitted regression line example

Sum of Squares:
Total: $SS_{tot} = \sum (y_i - \overline{y})^2$
Regression: $SS_{reg} = \sum (\hat{y}_i - \overline{y})^2$
Error: $SS_{err} = \sum (y_i - \hat{y}_i)^2$

Relationship: $SS_{tot} = SS_{reg} + SS_{err}$

Mean squared error: $MSE = \frac{SS_{err}}{n-2}$
Standard error: $s = \sqrt{MSE}$

Coefficient of Determination:

$$R^2 = \frac{SS_{reg}}{SS_{tot}} = 1 - \frac{SS_{err}}{SS_{tot}}$$

Interpretation: Proportion of variance explained by regression
Range: $0 \leq R^2 \leq 1$

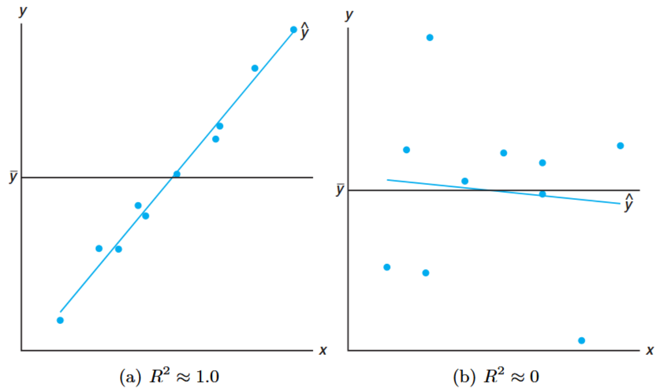

Figure 10: Good fit vs poor fit comparison in regression

Correlation Coefficient:
Sample correlation:

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}}$$

Properties:

- $-1 \leq r \leq 1$
- $r^2 = R^2$ in simple regression
- Sign of r matches sign of $\hat{\beta}_1$

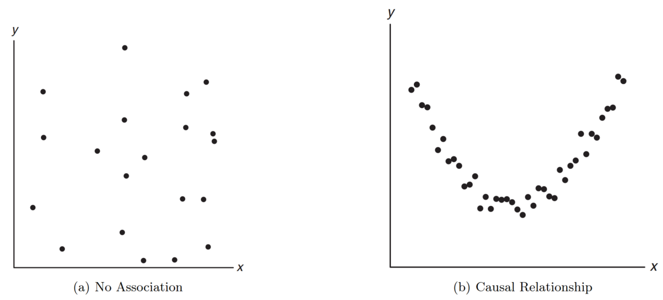

Figure 11: Examples of different correlation coefficients

Inference for Regression:
Standard errors:

$$SE(\hat{\beta}_1) = \frac{s}{\sqrt{S_{xx}}}, \quad SE(\hat{\beta}_0) = s\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{S_{xx}}}$$

Tests for slope:

$$T = \frac{\hat{\beta}_1 - \beta_1}{SE(\hat{\beta}_1)} \sim t_{n-2}$$

CI for slope: $\hat{\beta}_1 \pm t_{\alpha/2,n-2} \cdot SE(\hat{\beta}_1)$

Prediction interval for Y at x_0:

$$\hat{y}_0 \pm t_{\alpha/2,n-2} \cdot s\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \overline{x})^2}{S_{xx}}}$$

10 Common Critical Values

Standard Normal (Z):

Confidence	90%	95%	98%	99%
$z_{\alpha/2}$	1.64	1.96	2.33	2.58

One-tailed:

α	0.10	0.05	0.025	0.01
z_α	1.28	1.64	1.96	2.33

t-Distribution (Selected Values):

df	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.96	9.92
5	2.02	2.57	3.36	4.03
10	1.81	2.23	2.76	3.17
20	1.72	2.09	2.53	2.85
30	1.70	2.04	2.46	2.75
∞	1.64	1.96	2.33	2.58

11 Paired Tests and Two-Sample Tests

Paired t-Test:
For dependent samples (before/after, matched pairs):

$$H_0 : \mu_D = 0 \quad \text{vs} \quad H_1 : \mu_D \neq 0$$

Test statistic: $T = \frac{\overline{D}-0}{S_D/\sqrt{n}} \sim t_{n-1}$
where $D_i = X_i - Y_i$, $\overline{D} = \frac{1}{n}\sum D_i$

CI for μ_D: $\overline{d} \pm t_{\alpha/2,n-1}\frac{s_d}{\sqrt{n}}$

Example: Testing drug effectiveness on same patients

Pooled t-Test (Equal Variances):
For independent samples when $\sigma_1^2 = \sigma_2^2$:

$$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 \neq \mu_2$$

Test statistic:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

Pooled standard deviation:

$$S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$$

Use when: $\frac{s_1^2}{s_2^2}$ is close to 1

Welch's t-Test (Unequal Variances):
For independent samples when $\sigma_1^2 \neq \sigma_2^2$:
Test statistic:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Degrees of freedom:

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1-1} + \frac{(S_2^2/n_2)^2}{n_2-1}}$$

Use when variances are clearly different

12 Chi-Square Tests

Chi-Square Goodness of Fit:
Tests if sample follows specified distribution:

$$H_0 : \text{Data follows specified distribution}$$

Test statistic: $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$
where O_i = observed frequency, E_i = expected frequency

Degrees of freedom: $\nu = k - 1 - (\text{parameters estimated})$

Requirements: $E_i \geq 5$ for all categories

Example: Testing if dice is fair

Chi-Square Test of Independence:
Tests independence between two categorical variables:

$$H_0 : \text{Variables are independent}$$

Test statistic: $\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$

Expected frequency: $E_{ij} = \frac{\text{(Row i total)(Column j total)}}{\text{Grand total}}$

Degrees of freedom: $\nu = (r-1)(c-1)$
where r = rows, c = columns

Example: Testing if treatment and outcome are independent

Chi-Square Test for Homogeneity:
Tests if several populations have same proportions:

$$H_0 : p_{11} = p_{21} = \cdots = p_{k1}$$

Same formula as independence test
Different interpretation: comparing populations

Example: Comparing cure rates across hospitals

13 Power and Sample Size

Power of a Test:
Power $= 1 - \beta =$ Probability of rejecting false H_0

Factors affecting power:

- Larger effect size $\Rightarrow$ higher power
- Larger sample size $\Rightarrow$ higher power
- Larger $\alpha \Rightarrow$ higher power
- Smaller $\sigma \Rightarrow$ higher power

Power curve: Plot of power vs true parameter value

Sample Size Determination:
For testing $H_0 : \mu = \mu_0$ vs $H_1 : \mu = \mu_1$:
To achieve power $1 - \beta$ at significance α:

$$n = \left(\frac{(z_\alpha + z_\beta)\sigma}{\mu_1 - \mu_0}\right)^2$$

For two-sample tests:

$$n = \frac{2(z_\alpha + z_\beta)^2\sigma^2}{(\mu_1 - \mu_2)^2}$$

(per group)
For proportions:

$$n = \frac{(z_\alpha + z_\beta)^2[p_0(1-p_0) + p_1(1-p_1)]}{(p_1 - p_0)^2}$$

14 Decision Rules and Common Scenarios

When to Use Which Test:
One Sample:

- σ known, any n: Z-test
- σ unknown, $n < 30$: t-test (assume normality)
- σ unknown, $n \geq 30$: t-test or Z-test

Two Samples:

- Dependent samples: Paired t-test
- Independent, equal variances: Pooled t-test
- Independent, unequal variances: Welch's t-test
- Large samples: Z-test for proportions

Quality Control Example:
Production line: target diameter 10mm, tolerance ± 0.2mm
Sample 16 parts: $\overline{x} = 10.15$, $s = 0.18$

Test if process is on target ($H_0 : \mu = 10$ vs $H_1 : \mu \neq 10$):

$$t = \frac{10.15 - 10}{0.18/\sqrt{16}} = \frac{0.15}{0.045} = 3.33$$

With $t_{0.025,15} = 2.13$, reject H_0. Process needs adjustment.

Medical Trial Example:
New drug vs placebo for blood pressure reduction:
Drug: $n_1 = 30$, $\overline{x}_1 = 12.5$, $s_1 = 4.2$
Placebo: $n_2 = 30$, $\overline{x}_2 = 8.1$, $s_2 = 3.8$

Equal variances assumed:

$$s_p = \sqrt{\frac{29(4.2)^2 + 29(3.8)^2}{58}} = 4.0$$

$$t = \frac{12.5 - 8.1}{4.0\sqrt{\frac{1}{30} + \frac{1}{30}}} = \frac{4.4}{1.03} = 4.27$$

Highly significant: drug is effective

Survey Sampling Example:
Poll: 400 voters, 220 support proposition
95% CI for true proportion:

$$\hat{p} = \frac{220}{400} = 0.55$$

$$0.55 \pm 1.96\sqrt{\frac{0.55 \times 0.45}{400}} = 0.55 \pm 0.049 = (0.501, 0.599)$$

Margin of error: $\pm 4.9\%$

Reliability Testing Example:
Component lifetime follows exponential distribution.
Test 20 components, observe failures at times:
5, 12, 18, 25, 30, 35, 42, 48, 55, 62 hours

Maximum likelihood estimate: $\hat{\lambda} = \frac{n}{\sum t_i} = \frac{10}{332} = 0.030$

Estimated mean lifetime: $\hat{\mu} = \frac{1}{0.030} = 33.2$ hours

15 Probability Calculation Examples

Bayes' Theorem Example:
Disease prevalence: 1%. Test accuracy: 95% (both sensitivity and specificity)
Person tests positive. What's P(actually has disease)?
Let D = has disease, T^+ = tests positive

$$P(D|T^+) = \frac{P(T^+|D)P(D)}{P(T^+|D)P(D) + P(T^+|D^c)P(D^c)}$$
$$= \frac{0.95 \times 0.01}{0.95 \times 0.01 + 0.05 \times 0.99} = \frac{0.0095}{0.0590} = 0.161$$

Only 16.1% chance of actually having disease!

Law of Total Probability Example:
Three machines produce parts: A (50%), B (30%), C (20%)
Defective rates: A (2%), B (3%), C (5%)

Overall defective rate:

$P(\text{defective}) = 0.5 \times 0.02 + 0.3 \times 0.03 + 0.2 \times 0.05 = 0.029$

If part is defective, P(from machine A):

$$P(A|\text{defective}) = \frac{0.02 \times 0.5}{0.029} = 0.345$$

Combination Example:
Committee of 5 from 12 people (7 men, 5 women)
Find P(exactly 3 women):

$$P(X = 3) = \frac{\binom{5}{3}\binom{7}{2}}{\binom{12}{5}} = \frac{10 \times 21}{792} = \frac{210}{792} = 0.265$$

16 Useful Identities

Summation Identities:

$$\begin{aligned}
&\sum_{i=1}^{n}(x_i - \overline{x}) = 0 \\
&\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n} x_i^2 - n\overline{x}^2 \\
&\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n} x_i y_i - n\overline{xy}
\end{aligned}$$

Probability Rules:

$$\begin{aligned}
&P(A^c) = 1 - P(A) \\
&P(A \cup B) = P(A) + P(B) - P(A \cap B) \\
&P(A|B) = \frac{P(A \cap B)}{P(B)} \\
&P(A \cap B) = P(A|B)P(B)
\end{aligned}$$

Normal Approximations:
Binomial to Normal (continuity correction):

$$P(X = k) \approx P(k - 0.5 < Y < k + 0.5)$$

where $Y \sim N(np, np(1 - p))$

Poisson to Normal: $\lambda > 5$
$X \sim \text{Poisson}(\lambda) \approx N(\lambda, \lambda)$

Rule of thumb for normal approximation to binomial:
Use when $np \geq 5$ and $nq \geq 5$

Common Probability Relationships:
For independent events: $P(A \cap B) = P(A)P(B)$

Multiplication rule: $P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B)$

Complement rule: $P(A^c) = 1 - P(A)$

At least one: $P(\text{at least one success}) = 1 - P(\text{all failures})$

17 Distribution Relationships

Special Cases and Limits:
Hypergeometric → Binomial: As $N \to \infty$, $K/N \to p$

Binomial → Poisson: As $n \to \infty$, $p \to 0$, $np = \lambda$

Poisson → Normal: As $\lambda \to \infty$

t-distribution → Normal: As $\nu \to \infty$

Chi-square: $\chi^2_1 = Z^2$ where $Z \sim N(0, 1)$

F-distribution: $F_{1,\nu} = t^2_\nu$