Chapters 2–4: Probability, Random Variables, Distributions, Expectation

Operations with Sets

- $A \cap \emptyset = \emptyset$
- $A \cup \emptyset = A$
- $A \cap A' = \emptyset$
- $A \cup A' = S$
- $S' = \emptyset$
- $\emptyset' = S$
- $(A')' = A$
- $(A \cap B)' = A' \cup B'$
- $(A \cup B)' = A' \cap B'$

Permutation (Order Matters)

- With Repetition: n^r
- Without Repetition: $\boxed{{}_nP_r = \frac{n!}{(n-r)!}}$

Combinations (order doesn't matter)

- With Repetition: $\frac{(r+n-1)!}{r!\,(n-1)!}$
- Without Repetition: $\boxed{{}_nC_r = \frac{n!}{r!(n-r)!}}$

Circular Arrangement

$(n-1)!$

- Permutation with identical items: $\binom{n}{n_1, n_2, \dots, n_m} = \frac{n!}{n_1! n_2! \cdots n_m!}$
- Partition: $\binom{n}{n_1, n_2, \dots, n_m} = \frac{n!}{n_1! n_2! \cdots n_m!}$

- **Additive Rule** → $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- **Conditional Probability** → $P(B|A) = \frac{P(A \cap B)}{P(A)}$, provided $P(A) > 0$
- **Product Rule** → $P(A \cap B) = P(A)P(B|A)$, provided $P(A) > 0$
- **Independence of Events** → $P(A|B) = P(A)$ or $P(B|A) = P(B)$
  which would mean that $P(A \cap B) = P(A)P(B)$
- **Bayes' Rule** → $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$
- **Mutually Exclusive iff** → $P(A \cup B) = P(A) + P(B)$

- **Total Probability Theorem**

$P(A) = \sum_{i=1}^{k} P(A \cap B_i)$, for the partition $B_1, \dots, B_k$

$= \sum_{i=1}^{k} P(A|B_i)P(B_i)$, $B_i \cap B_u = \emptyset$, $B_i \cup B_k = S$

- **Bayes Rule with Total Probability**

$P(B|A) = \frac{P(B)P(A|B)}{\sum_{i=1}^{k} P(C_i)P(A|C_i)}$

for the partition $C_1, \dots, C_k$

	Discrete RV	Continuous RV
Probability Density	Probability Mass Function (PMF) • $f(x) \geq 0$, for each outcome $X = x$ • $\sum_x f(x) = 1$ •	Probability Density Function (PDF) • $f(x) \geq 0$ for each possible value $X = x$ • $\int_{-\infty}^{\infty} f(x)\,dx = 1$ • $\int_a^b f(x)\,dx = P(a < x < b)$
Cumulative Density Function (CDF)	• $P(X \leq x) = F(x) = \sum_{t \leq x} f(t)$ for $x \in \mathbb{R}$	• $P(X \leq x) = F(x) = \int_{-\infty}^{x} f(t)\,dt$ for $x \in \mathbb{R}$ • $P(a < X < b) = F(b) - F(a) = \int_a^b f(t)\,dt$
Joint Probability Distribution	Joint PMF • $f(x,y) \geq 0 \quad \forall (x,y) \in S$ • $\sum_x \sum_y f(x,y) = 1$ • $P(X=x, Y=y) = f(x,y)$ • $P((X,Y) \in A) = \sum_{(x,y)\in A} f(x,y)$	Joint PDF • $f(x,y) \geq 0 \quad \forall (x,y) \in S$ • $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$ • $P((X,Y) \in A) = \int_{(x,y)\in A} f(x,y)\,dx\,dy$
Marginal Distribution	• $g(x) = \sum_y f(x,y)$ • $h(y) = \sum_x f(x,y)$	• $g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$ • $h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$
Conditional Distributions	$P(a \leq X \leq b \mid Y=y) = \sum_{a \leq x \leq b} f(x \mid y)$	$P(a \leq x \leq b \mid Y=c) = \int_a^b f(x \mid Y=c)\,dx$
Expectation of a Function of	One RV $E[z(X)] = \sum_x z(x) f(x)$ Two RV $E[z(X,Y)] = \sum_y \sum_x z(x,y) f(x,y)$	One RV: $E[z(X)] = \int_{-\infty}^{\infty} z(x) f(x)\,dx$ Two RV $E[z(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} z(x,y) f(x,y)\,dx\,dy$
Variance of an RV	$\sigma^2 = Var(X) = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x)$ $= E(X^2) - \mu^2$	$\sigma^2 = Var(X) = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$ $= E(X^2) - \mu^2$
Covariance of RVs	$\sigma_{xy} = Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$ $= \sum_x \sum_y (x-\mu_x)(y-\mu_y) f(x,y)$ $= E[XY] - \mu_x \mu_y$	$\sigma_{xy} = Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$ $= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y) f(x,y)\,dx\,dy$ $= E[XY] - \mu_x \mu_y$

Conditional Distributions

$f(x|y) = \frac{f(x,y)}{g(y)}$

Independence of RVs

$f(x,y) = g(x)h(y)$

Expectation of two RVs

$E(aX + bY) = aE[X] + bE(Y)$ (any case)

$E(XY) = E[X]E[Y]$ (only if X and Y are independent)

Variance of a RV (any case)

$\sigma^2_{aX+bY+c} = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}$

Covariance of a RV any case

$\sigma_{XY} = E[XY] - E[X]E[Y]$

Correlation Coefficient of RVs

$-1 \leq \rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \leq 1$

Poisson Approximation for Binomial

$b(x; n, p) \to P(x; \lambda t)$ as $n \to \infty$, $p \to 0$ (np constant)

Chebyshev's Theorem (for discrete or continuous RVs)

$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$

Chapter 5 – Discrete Probability Distribution

Distribution	Probability Mass Function (PMF)	Expectation	Variance
Binomial	$b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$	$\mu = np$	$\sigma^2 = np(1-p)$
Multinomial	$f(x_1, \dots, x_m; p_1, \dots, p_m, n) = \binom{n}{x_1, \dots, x_m} p_1^{x_1} \cdots p_m^{x_m}$	$\mu_i = np_i$	$\sigma_i^2 = np_i(1-p_i)$
Hyper-Geometric	$h(x; N, n, K) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$ Approximates to ~Binomial iff $\frac{n}{N} < 0.05$	$\mu = \frac{nK}{N}$	$\sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{K}{n}\left(1 - \frac{K}{N}\right)$

Negative Binomial	$b^*(x;k,p)=\binom{x-1}{k-1}p^k q^{x-k}$ for $x\geq k$	$\mu=\frac{k}{p}$	$\sigma^2=\frac{k(1-p)}{p^2}$
Geometric	$g(x;p)=p(1-p)^{x-1}$ for $x\geq 1$	$\mu=\frac{1}{p}$	$\sigma^2=\frac{1-p}{p^2}$
Poisson	$p(x;\lambda t)=\frac{e^{-\lambda t}(\lambda t)^x}{x!}$ for $x\geq 0$	$\mu=\lambda t$	$\sigma^2=\lambda t$

Chapter 6 – Continuous Probability Distributions

Distribution	Probability Density Function (PDF)	Expectation	Variance
Uniform	$f(x;A,B)=\frac{1}{B-A}$, $A\leq x\leq B$	$\mu=\frac{A+B}{2}$	$\sigma^2=\frac{(B-A)^2}{12}$
Normal	$n(x;\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty<x<\infty$	μ	σ^2
Standard Normal CDF	$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$; $P(A\leq X\leq B)=\Phi(B)-\Phi(A)$	$\mu=0$	$\sigma^2=1$
Gamma	$f(x;\alpha,\beta)=\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}$, $x>0$	$\mu=\alpha\beta$	$\sigma^2=\alpha\beta^2$
Exponential	$f(x;\alpha=1,\beta)=f(x;\beta)=\frac{1}{\beta}e^{-\frac{x}{\beta}}$; $x\geq 0$, $\beta=\frac{1}{\lambda}$	$\mu=\beta$	$\mu=\beta^2$
Chi-Squared	$f(x;\alpha=\frac{v}{2},\beta=2)=f(x;v)=\frac{1}{2^{\frac{v}{2}}\Gamma(\frac{v}{2})}x^{\frac{v}{2}-1}e^{-\frac{x}{2}}$ for $x>0$	$\mu=v$	$\sigma^2=2v$

Standardized Variable

- $Z=\frac{X-\mu}{\sigma}$ (Z-score)
- $P(Z\leq\frac{x-\mu}{\sigma})=\int_{-\infty}^{\frac{x-\mu}{\sigma}}n(s;0,1)\,ds$
- $P(\frac{A-\mu}{\sigma}\leq Z\leq\frac{B-\mu}{\sigma})=\Phi(\frac{B-\mu}{\sigma})-\Phi(\frac{A-\mu}{\sigma})$
- $n(x;\mu,\sigma)=n(\frac{x-\mu}{\sigma};0,1)/\sigma$

Normal Approximation of Binomial Distribution

- $P(X\leq x)\approx P(Z\leq\frac{x+0.5-np}{\sqrt{npq}})$ for $np, n(1-p)\geq 5$
- $b(x;n,p)\approx n(x;np,\sqrt{npq})$

Poisson and Exponential Distribution

$\frac{d}{dx}P(X\leq x)=\lambda e^{-\lambda x}$, $\lambda=\frac{1}{\beta}$

Gamma Function (Properties)

$\Gamma(a)=\int_0^\infty x^{\alpha-1}e^{-x}dx, \alpha>0$ | $\Gamma(\frac{1}{2})=\sqrt{\pi}$ | $\Gamma(n)=(n-1)!, n\in\mathbb{N}$ | $\Gamma(1)=1$

Chapter 7 – Foundations of Random Variables

Transformation of Random Variables (One-to-one) → Bijective

- X (discrete) with PMF $f(x)$
  Let $Y=u(X)\longrightarrow y=u(x)$
  $x=w(y)=u^{-1}(y)$; $g(y)=f(w(y))$
- X (discrete) with Joint PMF $f(x_1,x_2)$
  Let $Y_1=U_1(X_1,X_2)$ & $Y_2=U_2(X_1,X_2)$
  $\longrightarrow y_1=u_1(x_1,x_2)$ & $y_2=u_2(x_1,x_2)$
  $x_1=w_1(y_1,y_2)$ and $x_2=w_2(y_1,y_2)$
  $g(y_1,y_2)=f[w_1(y_1,y_2),w_2(y_1,y_2)]$
- X (Continuous) with PDF $f(x)$
  Let $Y=u(X)\longrightarrow y=u(x)$
  $x=w(y)=u^{-1}(y)$; $f[w(y)]|J|$
  $J=\frac{dx}{dy}=w'(y)$
- X (Continuous) with Joint PDF $f(x_1,x_2)$
  Let $Y_1=u_1(X_1,X_2)$ & $Y_2=u_2(X_1,X_2)$
  $\longrightarrow y_1=u_1(x_1,x_2)$ and $y_2=u_2(x_1,x_2)$
  $x_1=w_1(y_1,y_2)$ and $x_2=w_2(y_1,y_2)$
  $g(y_1,y_2)=f[w_1(y_1,y_2),w_2(y_1,y_2)]|J|$
  $J=\begin{bmatrix}\frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2}\\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2}\end{bmatrix}$ 2 x 2

(Not one-to-one) (Non-Bijective) $Y=u(X)$ (Continuous)

$x_1=w_1(y), x_2=w_2(y)$ … $x_k=w_k(y)$

$g(y)=\sum_{i=1}^{k}f[w_i(y)]|J_i|$

where $J_i=w_i'(y), i=1,2,\dots k$

Moments and Moment-Generating Functions

- Uniqueness Theorem: $M_X(t)=M_Y(t)\rightarrow$ Same PDF
- r^{th} Moment about the origin
  $\mu_r'=E(X^r)=\begin{cases}\sum_x x^r f(x) & \text{(discrete)}\\ \int_{-\infty}^{\infty}x^r f(x)\,dx & \text{(continuous)}\end{cases}$
- Moment Generating Function
  $M_X(t)=E(e^{tX})=\begin{cases}\sum_x e^{tx}f(x) & \text{(discrete)}\\ \int_{-\infty}^{\infty}e^{tx}f(x)\,dx & \text{(continuous)}\end{cases}$
- $\boxed{\left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0}=\mu_r'}$
- $M_X(t)=e^{\mu t+(\frac{t^2\sigma^2}{2})}$ Normal Distribution

Linear Combination of Random Variables

For $Y=aX$ given distribution $f(x)$
$\longrightarrow h(y)=\frac{1}{|a|}f(\frac{y}{a})$

- $M_{X+a}(t)=e^{at}M_X(t)$
- $M_{aX}(t)=M_X(at)$

- For $Z=X+Y$ and distributions $f(x), g(y)$
  X and Y independent, $X=W$, $Y=Z-W$
  $h(z)=\sum_{-\infty}^{\infty}f(w)g(z-w)$ (Discrete)
  $h(z)=\int_{-\infty}^{\infty}f(w)g(z-w)\,dw$ (Continuous)

$X_1, X_2, \dots X_n$ are independent w/ MGF
$M_{X_1}(t), M_{X_2}(t), \dots M_{X_n}(t)$; $Y=X_1+X_2+\dots+X_n$
$\therefore M_Y(t)=M_{X_1}(t)M_{X_2}(t)\cdots M_{X_n}(t)$

$Y=a_1X_1+a_2X_2+\dots+a_nX_n$
$\mu_Y=a_1\mu_1+\dots+a_n\mu_n$; $\sigma_Y^2=a_1^2\sigma_1^2+\dots+a_n^2\sigma_n^2$

Chapter 8 – Fundamental Sampling Distributions and Data Description

Sample Random Variables (Statistics)

Sample Data $x_1, x_2, \dots, x_n$
Each independent measures of RV X_i

- $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ (empirical value of mean)
- $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ (Sample mean RV)

Standard Deviation & Mean

- $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})$ (Sample variance)

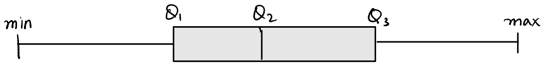
Mean, Median Mode, Range

Data in increasing order $x_{(1)}, \dots, x_{(n)}$

- Median $=\frac{x_{(\frac{n}{2})}+x_{(\frac{n}{2}+1)}}{2}$, n is even ; $x_{(\frac{n+1}{2})}$, n is odd
- Range $=\max(x_i)-\min(x_i)$
- $S=\sqrt{S^2}$ (Sample Standard Deviation)
- $E(\bar{X})=\mu$ and $E(S^2)=\sigma^2$ (unbiased)
- $S^2=\frac{1}{n(n-1)}\left[n\sum_{i=1}^{n}X_i^2-\left(\sum_{i=1}^{n}X_i\right)^2\right]$

Box-and-Whisker

- $Q_i = (n+1)\cdot\frac{i}{4}$; $i = 1, 2, 3$ (Quartiles)
- Interquartile Range (IQR) $= Q_3 - Q_1$
- Q_2 is the median
- Lower Whisker (minimum): $Q_1 - 1.5\,(IQR)$
- Upper Whisker (maximum): $Q_3 - 1.5\,(IQR)$



Quantile Plots

- Quantile Plot: $\left(\dfrac{i-\frac{3}{8}}{n+\frac{1}{4}}, x_i\right) = (f, x_i)$
- $q_{\mu,\sigma}(f) = \mu + \sigma\left[4.91\left[f^{0.14} - (1-f)^{0.14}\right]\right]$
- (Normal Q-Q) $= (q_{0,1}(f_i), x_i)$
  $q_{0,1} = 4.9\left[f^{0.14} - (1-f)^{0.14}\right]$
  For CDF $F(x)$, $q(f) = F^{-1}$

Chapter 9 – One and Two-Sample Estimation Problems

Normal Distribution Facts • X_1, X_2 independent normal RVs, $X_1 + X_2$ normal, $\mu = \mu_1 + \mu_2$, and $\sigma^2 = \sigma_1^2 + \sigma_2^2$

If X normal, then $\frac{X}{n}$ normal, $\frac{\mu}{n}$, $\frac{\sigma^2}{n^2}$; $x_1, \dots, x_n$ independent normal, $\bar{X}$ normal, μ, σ^2/n

Central Limit Theorem	T-distribution	Chi-Squared (Variance)
$Z_n = \dfrac{\bar{X}_n - \mu}{\sigma}$ as $n \to \infty$ $Z_n \to n(z; 0, 1)$, $n \geq 30$	$T = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$ $n < 30$, $v = n-1$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$ $v = n-1$

Point Estimates
$\theta = \mu$, $\hat{\theta} = \bar{x}$, $\hat{\theta} = \bar{X}$
$E(\hat{\theta}) = \theta$ (unbiased estimator)

Confidence Intervals Table

Purpose	$P(\Theta_L < \theta < \Theta_U) = 1-\alpha$
Mean (known σ & $n \geq 30$)	$P\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$
Mean (known σ, $n<30$) (unknown σ)	$P\left(\bar{X} - t_{\alpha/2}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 1-\alpha$
Prediction Intervals	For the next observation x_0: $P\left(\bar{x} - z_{\alpha/2}\,\sigma\sqrt{1+\frac{1}{n}} \leq x_0 \leq \bar{x} + z_{\alpha/2}\,\sigma\sqrt{1+\frac{1}{n}}\right) = 1-\alpha$
Difference of Means	Known Population Variance (σ_1 and σ_2) $P\left((\bar{x}_1-\bar{x}_2) - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \leq \mu_1-\mu_2 \leq (\bar{x}_1-\bar{x}_2) - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$ $= 1-\alpha$
	Unknown and Equal Population Variances ($\sigma_1 = \sigma_2$) $S_p^2 = \dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$, $n_1, n_2 < 30$, $v = n_1+n_2-2$ $P\left((\bar{x}_1-\bar{x}_2) - t_{\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \leq \mu_1-\mu_2 \leq (\bar{x}_1-\bar{x}_2) + t_{\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right) = 1-\alpha$
	Unknown and Unequal Population Variances ($\sigma_1 \neq \sigma_2$) $v = \left[\dfrac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}}\right]$, $P\left((\bar{x}_1-\bar{x}_2) - t_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} \leq \mu_1-\mu_2 \leq (\bar{x}_1-\bar{x}_2) + t_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}\right) = 1-\alpha$
Paired Observation	$D_i = X_{1,i} - X_{2,i}$, $\mu_D = \mu_1 - \mu_2$, $\bar{d} = \bar{x}_1 - \bar{x}_2$, $v = n-1$ variance$(D_i) = \sigma^2_{X_{1,i}} + \sigma^2_{X_{2,i}} - 2\,$Covariance$(X_{1,i}, X_{2,i})$ $P\left(\bar{d} - t_{\frac{\alpha}{2}}\frac{S_d}{\sqrt{n}} \leq \mu_D \leq \bar{d} + t_{\frac{\alpha}{2}}\frac{S_d}{\sqrt{n}}\right) = 1-\alpha$
Estimating a Proportion (Single Sample)	$\hat{p} = \frac{X}{n}$, $\hat{p} = \frac{x}{n}$, p unknown (Binomial Distribution) $P\left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 1-\alpha$ $n = \dfrac{z^2_{\frac{\alpha}{2}}\,\hat{p}(1-\hat{p})}{\delta^2}$ $\quad n \geq \dfrac{z^2_{\frac{\alpha}{2}}}{4\delta^2}$, $\max(\hat{p}(1-\hat{p})) = 0.25$
Variance (σ^2)	Use Chi-Squared Distribution $P\left(\dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} \leq \sigma^2 \leq \dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}}\right) = 1-\alpha$, $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$ $v = n-1$ (degrees of freedom)

For Upper and Lower Bounds

- $P(\theta \leq \Theta_U) = 1-\alpha$ (upper bound)
- $P(\theta \geq \Theta_L) = 1-\alpha$ (lower bound)
  - $P\left(\mu \leq \bar{X} + z_\alpha\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$ (upper bound)
  - $P\left(\mu \geq \bar{X} - z_\alpha\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$ (lower bound)

standard error $= \dfrac{\sigma}{\sqrt{n}}$; margin error $= z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$

Tolerance Limits (Tolerance Factor Table)

$P(\bar{x} \pm ks) = 1-\gamma$ (that $1-\alpha$ of samples in range)

Maximum Likelihood Estimation and Log Likelihood

- Samples $x_1, x_2, \dots, x_n$ with Joint Probability Density Function $f(x_1, x_2, \dots, x_n; \theta)$
- $L(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} g(x_i; \theta)$

Maximum Likelihood

- $\hat{\theta} = \arg\max L(x_1, \dots, x_n; \theta) = \theta$
  Such that $\dfrac{dL}{d\theta} = 0$ and $\dfrac{d^2L}{d\theta^2} < 0$

Log-Likelihood

- $\text{Log } L = \text{Log}\left(\prod_{i=1}^{n} g(x_i; \theta)\right)$
- $\hat{\theta} = \arg\max\left[\text{Log } L(x_1, \dots, x_n; \theta)\right] = \theta$

  such that $\dfrac{d(\text{Log } L)}{d\theta} = 0$

- $\mu = \dfrac{\partial}{\partial\mu}\ln\left(L(x_1, \dots, x_n; \mu, \sigma)\right) = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
- $\sigma^2 = \dfrac{\partial}{\partial\sigma^2}\ln\left(L(x_1, \dots, x_n; \mu, \sigma)\right)$
  $= \dfrac{n-1}{n}S^2 = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Chapter 10 - One and Two Sample Tests of Hypothesis

- Type 1 (Error) $\alpha = P_r(\text{Reject } H_0 \mid H_0 \text{ is True}) = P(\text{Critical Region})$ with H_0
- Type 2 (Error) $\beta = P_r(\text{Fail to reject } H_0 \mid H_0 \text{ is False}) = 1 - P(\text{Critical Region})$ with H_1

Decision	H_0 = True	H_0 = False
Fail to reject H_0	Correct	Type II
Reject H_0	Type I	Correct

Table: Hypothesis Tests (Concerning Means) [Reject if]

H_0	Value of Test Statistic	H_1	Critical Region
$\mu = \mu_0$	$z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$; σ known	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$z < -z_\alpha$ $z > -z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu = \mu_0$	$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$; $v = n-1$ σ unknown	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$z = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ σ_1 and σ_2 known	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $v = n_1 + n_2 - 2$ $\sigma_1 = \sigma_2$ but known $s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t' = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $v = \dfrac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{(n_1-1)} + \frac{(s_2^2/n_2)^2}{(n_2-1)}}$ $\sigma_1 \neq \sigma_2$ and unknown	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t' < -t_\alpha$ $t' > t_\alpha$ $t' < -t_{\alpha/2}$ or $t' > t_{\alpha/2}$
$\mu_D = d_0$ (paired observation)	$t = \dfrac{\bar{d} - d_0}{s_d/\sqrt{n}}$, $v = n-1$	$\mu_D < d_0$ $\mu_D > d_0$ $\mu_D \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$

Table: Hypothesis Tests (Concerning Variances) [Reject if]

H_0	Test Statistic	H_1	Critical Region
One Variance $\sigma^2 = \sigma_0^2$	(Chi-Squared Distribution) $\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$, $v = n-1$	$\sigma^2 < \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha}$ $\chi^2 > \chi^2_\alpha$ $\chi^2 < \chi^2_{1-\frac{\alpha}{2}}$ or $\chi^2 > \chi^2_{\frac{\alpha}{2}}$
Two Variance $\sigma_1^2 = \sigma_2^2$	f-Distribution $f_{1-\alpha}(v_1, v_2) = \dfrac{1}{f_\alpha(v_2, v_1)}$ $f = \dfrac{s_1^2}{s_2^2}$, $v_1 = n_1 - 1$, $v_2 = n_2 - 1$	$\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 \neq \sigma_2^2$	$f < f_{1-\alpha}(v_1, v_2)$ $f > f_\alpha(v_1, v_2)$ $f < f_{1-\frac{\alpha}{2}}$ or $f > f_{\alpha/2}$

→ Two Tailed Tests

Statistical Power $= 1 - \beta$

Hypothesis Testing:

- Null H_0: $\theta = \theta_0$
- Alternate H_1: $\theta \neq \theta_0$ (two-sided)

One Sided $\begin{cases} \theta > \theta_0 \text{ (upper bound)} \\ \theta < \theta_0 \text{ (lower bound)} \end{cases}$

Approach 1 → P-value

- P(critical region) $< \alpha$
  Reject H_0
- P(critical region) $> \alpha$
  Fail to reject H_0
- P(Non-Critical Region) $> 1-\alpha$
  Reject H_0
- P(Non-Critical Region) $< 1-\alpha$
  Fail to Reject H_0

Approach 2 - Critical Region

$P(r_{\frac{\alpha}{2}} < r < r_{1-\frac{\alpha}{2}}) = 1 - \alpha$

for f and χ^2-distributions.

Goodness of Fit

- $e_i = nP(i)$, n trials, k outcomes
- $\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$, $v = k-1$
- Small χ^2 = Good Fit
- Reject if $\chi^2 > \chi^2_\alpha$

Linear Regression

- Data $\{(x_i, y_i);\ i = 1, 2, 3, \dots, n\}$
- Real: $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$,
  where $E(\epsilon) = 0$ and $Var(\epsilon)$
- Estimator: $\hat{y}_i = b_0 + b_1 x_i$
- Residual: $e_i = y_i - \hat{y}_i$; $i = 1, \dots, n$
- $y_i = b_0 + b_1 x_i + e_i$

Chapter 11 - Simple Linear Regression and Correlation

- $SSE = \sum_{i=1}^n e_i^2$, $\frac{\partial(SSE)}{\partial b_0} = 0$, $\frac{\partial(SSE)}{\partial b_1} = 0$
- $b_0 = \bar{y} - b_1\bar{x} = \frac{1}{n}\left(\sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i\right)$
- $b_1 = \dfrac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \dfrac{n\sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}$

Confidence Interval for Regression Parameters $\mu_{Y|x} = \beta_0 + \beta_1 x$

- $P\left(b_1 - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{S_{xx}}} \le \beta_1 \le b_1 + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{S_{xx}}}\right) = 1 - \alpha$, $v = n-2$
- $P\left(b_0 - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n S_{xx}}}\sqrt{\sum_{i=1}^n x_i^2} \le \beta_0 \le b_0 + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n S_{xx}}}\sqrt{\sum_{i=1}^n x_i^2}\right) = 1 - \alpha$, $v = n-2$

$t = \dfrac{b_1 - \beta_1}{s/\sqrt{S_{xx}}}$ (slope), $t = \dfrac{b_0 - \beta_0}{s\sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{xx}}}}$ (Intercept)

Sum of Errors

- $S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n}\left(\sum_{i=1} x_i\right)^2 = \sum_{i=1}^n x_i^2 - n(\bar{x})^2$
- $S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - \frac{1}{n}\left(\sum_{i=1}^n y_i\right)^2 = \sum_{i=1}^n y_i^2 - n(\bar{y})^2$
- $S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \frac{1}{n}\left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}$
- $SSE = S_{yy} - b_1 S_{xy} = S_{yy} - \left(\frac{S_{xy}}{S_{xx}}\right)S_{xy} = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$
- $s^2 = E[\sigma^2] = \frac{SSE}{n-2}$

Hypothesis Testing with Regression Parameters

$H_0: \beta_1 = \beta_{10}$ $H_1: \beta_1 \neq \beta_{10}$ $H_1: \beta_1 < \beta_{10}$ or $\beta_0 > \beta_{10}$

$H_0: \beta_1 = \beta_{00}$ $H_1: \beta_1 \neq \beta_{00}$ $H_1: \beta_1 < \beta_{00}$ or $\beta_0 > \beta_{00}$

$t = \dfrac{b_0 - \beta_{00}}{s\sqrt{\frac{\sum_{i=1}^n x_i^2}{n S_{xx}}}}$ Intercept ; $t = \dfrac{b_1 - \beta_{10}}{s/\sqrt{S_{xx}}}$

Coefficient of Determination (R^2) $\quad R^2 = 1 - \dfrac{SSE}{SST} = 1 - \sum_{i=1}^n \dfrac{(y_i - \hat{y}_i)^2}{(y_i - \bar{y}_i)^2}$, $0 \le R^2 \le 1$