A Bird's-eye View

Mathematical Statistics

Dane Jeon

Chapter 1 *Probability*

(def) *Prior Probabilities*

Posterior Probabilities

Probability Axioms

Sample Space Ω

Sample Point ω

σ*-field* $\mathcal{F}$

Probability Set Function $P : \mathcal{F} \to \mathbb{R}$

Probability Space $(\Omega, \mathcal{F}, P)$

(def) *Event*

Disjoint, Mutually Exclusive

(thm) *Inclusion-Exclusion Principle*

(def) *Conditional Probability*

(thm) *Multiplication Rule*

(def) *Partition*

Disjoint Union

Exhaustive

(thm) *Rule of Total Probability*

(thm) *Bayes' Theorem*

(def) *(Stochastic) Independence*

Mutual Independence

Chapter 2 *Random Variables*

(def) *Random Variable* X

Range, Space $\mathcal{D}$

Distribution

Probability Mass Function f_X

Probability Density Function f_X

Cumulative Distribution Function F_X

(def) *Discrete Random Variable* X

Support $\mathcal{S}$

Transformation

(thm)

$$f_Y(y) = P\left[Y = y\right] = P\left[g(X) = y\right] = P\left[X = g^{-1}(y)\right] = f_X(g^{-1}(y))$$

(def) *Continuous Random Variable* X

(thm) *Cumulative Distribution Function Technique*

(thm)

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dx}{dy} \right|$$

(def) *Jacobian* J

Mixture

(def) *Expectation* $E\left[X\right]$

(thm)

$$E\left[Y\right] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \qquad \text{and} \qquad E\left[Y\right] = \sum_{x \in \mathcal{S}_X} g(x) f_X(x)$$

(thm)

$$E\left[k_1 g_1(X) + k_2 g_2(X)\right] = k_1 E\left[g_1(X)\right] + k_2 E\left[g_2(X)\right]$$

(def) *Mean Value* μ or $E\left[X\right]$

Variance σ^2 or $\mathrm{Var}\left[X\right]$

Standard Deviation σ or $\mathrm{sd}\left[X\right]$

(def) *Moment Generating Function* $m_X(t)$

n*-th Moment* $E\left[X^n\right]$

(def) *Markov's Inequality*

Chebychev's Inequality

Chapter 3 *Joint Probability Distributions*

(def) *Random Vector* (X_1, X_2)

Range, Space $\mathcal{D}$ or R_X

Joint Cumulative Distribution Function F_{X_1,X_2}

Discrete Random Vector

Joint Probability Mass Function f_{X_1,X_2}

Continuous Random Vector

Joint Probability Density Function f_{X_1,X_2}

Support $\mathcal{S}$

Marginal Probability Mass Functions

Marginal Probability Density Functions

Conditional Probability Mass Functions $f_{X_1|X_2}(x_2|x_1)$

Conditional Probability Density Functions $f_{X_1|X_2}(x_2|x_1)$

(def) *(Stochastic) Independence*

(def) *Covariance* $\mathrm{Cov}[X_1, X_2]$

Correlation Coefficient $\rho[X_1, X_2]$

(thm) Let $T = \sum_{i=1}^{n} a_i X_i$. Then,

$$E[T] = \sum_{i=1}^{n} a_i E[X_i]$$

(thm) Let $T = \sum_{i=1}^{n} a_i X_i$ and $W = \sum_{j=1}^{m} a_i Y_i$. Then,

$$\mathrm{Cov}[T, W] = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j \mathrm{Cov}[X_i, Y_j]$$

(cor) Let $T = \sum_{i=1}^{n} a_i X_i$. Then,

$$\mathrm{Var}[T] = \mathrm{Cov}[T, T] = \sum_{i=1}^{n} a_i^2 \mathrm{Var}[X_i] + 2 \sum_{i<j} a_i a_j \mathrm{Cov}[X_i, X_j]$$

(cor) If $X_1, ..., X_n$ are independent random variables,

$$\mathrm{Var}[T] = \sum_{i=1}^{n} a_i^2 \mathrm{Var}[X_i]$$

Chapter 4 *Discrete Probability Distributions*

(def) *Discrete Uniform Distribution*

$$f_X(x) = \frac{1}{n} \cdot I_{R_X}$$

Bernoulli Distribution

$$f_X(x) = p^x(1-p)^{1-x} \cdot I_{R_X}$$

Binomial Distribution

$$f_X(x) = \binom{n}{x} p^x(1-p)^{n-x} \cdot I_{R_X}$$

Geometric Distribution

$$f_X(x) = (1-p)^{x-1}p \cdot I_{R_X}$$

Negative Binomial Distribution

$$f_X(x) = \binom{x-1}{r-1} p^r(1-p)^{x-r} \cdot I_{R_X}$$

Poisson Distribution

$$f_X(x) = \frac{m^x e^{-m}}{x!} \cdot I_{R_X}$$

(thm) Let $X \sim B(m, p)$ and $Y \sim B(n, p)$ be independent. Then,

$$X + Y \sim B(m + n, p)$$

(thm) *Memorylessness (of geometric distributions)*

(thm) Let $X_i \sim \mathrm{Ber}(p)$ and $X_1, ..., X_n$ be pairwise independent. Then,

$$X_1 + ... + X_r \sim \mathrm{B}(n, p)$$

(thm) Let $X_i \sim \mathrm{Ge}(p)$ and $X_1, ..., X_r$ be pairwise independent. Then,

$$X_1 + ... + X_r \sim \mathrm{NB}(r, p)$$

Chapter 5 *Continuous Probability Distributions*

(def) *Continuous Uniform Distribution*

$$f_X(x) = \frac{1}{(b-a)} \cdot I_{R_X}$$

Gamma Distribution

$$f_X(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha} \cdot I_{R_X}, \qquad \alpha > 0,\ \beta > 0$$

Exponential Distribution

$$f_X(x) = \lambda e^{-\lambda t} \cdot I_{R_X}$$

Normal Distribution

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \cdot I_{R_X}$$

Standard Normal Distribution

Standard Normal Random Variable

(def) *Gamma Function*

$$\Gamma(\alpha) = \int_{-\infty}^{\infty} t^{\alpha-1}e^{-t}dt, \qquad \alpha > 0$$

(thm) Let $X \sim P(m)$ and T be the time taken until the α-th event. Then, $T \sim \mathrm{Ga}(\alpha, 1/m)$

(thm) *Memorylessness (of exponential distributions)*

(def) *Survival Function* $S(x)$

Hazard Rate, Failure Rate $h(x)$

(thm) X has a $N(\mu, \sigma^2)$ distribution if and only if $Z = (X-\mu)/\sigma$ has a $N(0,1)$ distribution.

(thm) A linear transformation $Y = aX + b$ of a $X \sim N(\mu, \sigma^2)$ is $Y \sim N(a\mu + b, (a\sigma)^2)$

(thm) Let $X_1, ..., X_n$ be pairwise independent with $X_i \sim N(\mu_i, \sigma_i^2)$. Then,

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n}(a_i\sigma_i)^2\right)$$

Chapter 7 *Sampling Distribution*

7.1 *Population Distributions*

(def) *Population* X or $f(x;\theta)$

Sample $X_1, X_2, ..., X_n$

Realizations $x_1, ..., x_n$

Sample Size n

Population Parameter θ

Random Sample $X_1, X_2, ..., X_n$

Statistic $T(X_1, X_2, ..., X_n)$ (of a random variable)

7.2 *Sampling Distributions from a Single Random Sample*

(def) *Sample Mean* $\bar{X}$ (of a random sample)

Sample Variance S^2 (of a random sample)

Population Proportion $\hat{p}$ (of a random sample each from a Bernoulli distribution)

(thm) *Central Limit Theorem* The linear transformation $Z = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ converges in distribution to $N(0, 1)$. That is, $\bar{X}$ converges in distribution to $N(\mu, \sigma^2/n)$.

(cor) Let $X_1, ..., X_n$ be a random sample with $X_i \sim N(\mu, \sigma^2)$. Then, $(n-1)S^2/\sigma^2$ converges in distribution to $\mathcal{X}^2(n-1)$.

(cor) Let $X_1, ..., X_n$ be a random sample with $X_i \sim Ber(p)$. Then, $\hat{p}$ converges in distribution to $N(p, p(1-p)/n)$.

(cor) Let $X_1, ..., X_n$ be a random sample with $X_i \sim N(\mu, \sigma^2)$. Then, $T = (\bar{X} - \mu)/(S/\sqrt{n})$ converges in distribution to $t(n-1)$.

7.3 *Sampling Distributions from Multiple Random Samples*

It is important to recognize that throughout, we assume the random variables from the two random samples are pairwise independent.

(cor) Let $X_1, ..., X_m$ be a random sample with $X_i \sim N(\mu_1, \sigma_1^2)$ and $Y_1, ..., Y_n$ be a random sample with $Y_i \sim N(\mu_2, \sigma_2^2)$. Then, $\bar{X} - \bar{Y}$ converges in distribution to $N(\mu_1 - \mu_2, \sigma_1^2/m + \sigma_2^2/n)$.

(cor) Let $X_1, ..., X_m$ be a random sample with $X_i \sim \text{Ber}(p_1)$ and $Y_1, ..., Y_n$ be a random sample with $Y_i \sim \text{Ber}(p_2)$. Then, $\hat{p}_1 - \hat{p}_2$ converges in distribution to $N(p_1 - p_2, p_1(1-p_1)/m + p_2(1-p_2)/n)$

(cor) Let $X_1, ..., X_m$ be a random sample with $X_i \sim N(\mu_1, \sigma_1^2)$ and $Y_1, ..., Y_n$ be a random sample with $Y_i \sim N(\mu_2, \sigma_2^2)$. Then, $(S_1^2/\sigma_1^2)/(S_2^2/\sigma_2^2)$ converges in distribution to $F(m-1, n-1)$.

(cor) (T when $\sigma_1^2 = \sigma_2^2 = \sigma^2$) Let $X_1, ..., X_m$ be a random sample with $X_i \sim N(\mu_1, \sigma^2)$ and $Y_1, ..., Y_n$ be a random sample with $Y_i \sim N(\mu_2, \sigma^2)$. Then, $T = \left[(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right]/S_p\sqrt{1/m + 1/n}$ converges in distribution to $t(m+n-2)$.

(cor) (T when $\sigma_1^2 \neq \sigma_2^2$) Let $X_1, ..., X_m$ be a random sample with $X_i \sim N(\mu_1, \sigma_1^2)$ and $Y_1, ..., Y_n$ be a random sample with $Y_i \sim N(\mu_2, \sigma_2^2)$. Then, $T = \left[(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)\right] / \sqrt{S_1^2/m + S_2^2/n}$ converges in distribution to $t(\nu)$.

7.4 *Distributions Generated by Standard Normal Distributions*

($\mathcal{X}^2$*-distribution*)

(thm) Let $X_1, ..., X_n$ be a random sample with $X_i \sim N(0,1)$. Then, $\sum_{i=1}^n X_i^2$ converges in distribution to $\mathcal{X}^2(n)$.

(cor) Let $X_1, ..., X_n$ be a random sample with $X_i \sim N(\mu_i, \sigma_i^2)$. Then, $\sum_{i=1}^n \left[(X_i - \mu_i)/\sigma_i\right]^2$ converges in distribution to $\mathcal{X}^2(n)$.

(cor) Let $X_1, ..., X_n$ be a random sample with $X_i \sim \mathcal{X}_i(k_i)$. Then, $\sum_{i=1}^n \mathcal{X}_i(k_i)$ converges in distribution to $\mathcal{X}^2(k_1 + k_2 + ... + k_n)$.

(cor) Let X_1, X_2 be independent with $X_i \sim \mathcal{X}_i(k_i)$. Then, $X_2 - X_1 \sim \mathcal{X}_i(k_2 - k_1)$ given that $k_2 - k_1 > 0$.

(t*-distribution*)

(thm) Let W and V be independent with $W \sim N(0,1)$ and $V \sim \mathcal{X}^2(k)$. Then, $T = W/\sqrt{V/k}$ follows the distribution $t(k)$.

(thm) $t_\alpha(k) = -t_{1-\alpha}(k)$

(F*-distribution*)

(thm) Let U and V be independent with $U \sim \mathcal{X}^2(m)$ and $V \sim \mathcal{X}^2(n)$. Then, $F = (U/m)/(V/n)$ follows the distribution $F(m,n)$.

(thm) $F_\alpha(m,n) = 1/F_{1-\alpha}(n,m)$

(thm) Let W and V be independent with $W \sim N(0,1)$ and $V \sim \mathcal{X}^2(k)$. Then, $T^2 = W^2/(V/k)$ follows the distribution $F(1,k)$.

7.5 *Order Statistics*

(def) Order Statistic $X_{(k)}$

(thm)

$$F_{X_{(k)}} = \sum_{j=k}^{n} \binom{n}{j} \left[F(x)\right]^j \left[F(x)\right]^{n-j}$$

Chapter 8 *Estimation*

(def) *Statistical Inference*

Point Estimation

Interval Estimation

Degree of Confidence

Estimator $\hat{\theta} = \hat{\theta}(X_1, ..., X_n)$

Estimate $\hat{\theta} = \hat{\theta}(x_1, ..., x_n)$

(def) *Unbiased Estimator*

Efficient Estimator

Consistent Estimator

Likelihood Function $L(\theta; x_1, x_2, ..., x_n)$

Maximum Likelihood Estimator

Chapter 9 *Hypothesis Testing*

This chapter deals with the (1.1) basics of hypothesis testing and (1.2) testing of normal distributions.

(def) *Null Hypothesis* H_0

Alternative Hypothesis H_1

Significance Level α

Critical Region

One Sided Lower Hypothesis

One Sided Higher Hypothesis

Two sided Hypothesis

Type I Error

Type II Error

Test Statistic

References

[1] R.V. Hogg, J.W. McKean, and A.T. Craig. *Introduction to Mathematical Statistics.* What's New in Statistics Series. Pearson, 2019.