Honours Algebra
Quick Notes

Ian S.W. Ma

Spring 2020

Contents

1 Statistics Basics

General Structure: Probability Space/Triple

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

- Ω: **Sample Space**, set of all posible outcomes.
- $\mathcal{F}$: **Set of events** σ, each event is a subset of Ω
- $\mathbb{P}$ Probability of event $\sigma \in \mathcal{F}$

Conditional Probability

Conditional probability of A given B: $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \Leftrightarrow \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(A \cap B)$

General properties

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

  Events are mutually exclusive $\Rightarrow \mathbb{P}(A \cap B) = 0$
- $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$

  Events are independent $\Rightarrow \mathbb{P}(A|B) = \mathbb{P}(A)$
- Law of Total Probability

  If events A_1, A_2, A_N are mutually exclusive and exhaustive ($\bigcup_{i=1}^N A_i = \Omega$) then:

$$\mathbb{P}(B) = \sum_{i=1}^{N} \mathbb{P}(A_i \cap B) = \sum_{i=1}^{N} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

Random Variables

$X \sim$ Distribution	$\mathbb{E}(X)$	$\mathrm{Var}(X)$
Normal(μ, σ^2)	μ	σ^2
Bernoulli(θ)	θ	$\theta(1-\theta)$
Binomial(θ)	$n\theta$	$n\theta(1-\theta)$
Poisson(μ)	μ	μ
Uniform(α, β)	$\frac{\alpha+\beta}{2}$	$\frac{(\alpha+\beta)^2}{12}$
Beta(α, β)	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Gamma(α, β)	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
Inv. Gamma(α, β)	$\frac{\beta}{\alpha-1}$	$\frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$
Exponential(λ)	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Random Vectors

$\mathbf{X} = X_1, ..., X_k \sim$ Distribution	$\mathbb{E}(X_i)$	$\mathrm{Var}(X_i)$	$\mathrm{Cov}(X_i, X_j)$
Multivariative Normal$(\boldsymbol{\mu}, \Sigma)$	μ_i	$\Sigma_{i,i}$	$\Sigma_{i,j}$
Multinomial$(n, \theta_1, ..., \theta_k)$	$n\theta_i$	$n(1-\theta_i)$	$-n\theta_i\theta_j$

2 Bayes Throrem

Discrete Case

$$
\begin{aligned}
\mathbb{P}(X = x_i|Y = y_j) &= \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)} \\
&= \frac{\mathbb{P}(Y = y_j|X = x_i)\mathbb{P}(X = x_i)}{\mathbb{P}(Y = y_j)} \\
&= \frac{\text{Likelihood} \times \text{Prior}}{\text{Marginal}} \\
&= \frac{\mathbb{P}(Y = y_j|X = x_i)\mathbb{P}(X = x_i)}{\sum_{k=1}^{n} \mathbb{P}(X = x_k, Y = y_j)} \\
&= \frac{\mathbb{P}(Y = y_j|X = x_i)\mathbb{P}(X = x_i)}{\sum_{k=1}^{n} \mathbb{P}(Y = y_j|X = x_k)P(X = x_k)} \\
\Rightarrow \mathbb{P}((X|Y)) &= \frac{\mathbb{P}(Y|X)\mathbb{P}(X)}{\sum_x \mathbb{P}(X = x, Y)} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Marginal}}
\end{aligned}
$$

Continous Case

$$
f(X|Y) = \frac{f(X, Y)}{g(Y)} = \frac{g(Y|X)f(X)}{\int g(Y|X)f(X)dX}
$$

Bayesian Statistical Inference

- $L(\theta|y) = f(y|\theta)$: Likelihood
- $\pi(\theta)$: Prior
- $m(y) = p(y)$: Marginal distribution/Normalizing constant

$$
\text{Posterior Distribution: } p(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int p(\theta, y)d\theta} = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d\theta} \propto f(y|\theta)\pi(\theta)
$$

Predictive Distrubutions

$$
\begin{aligned}
\text{Prior Predictive Distribution:} \qquad & p(y^{new}) = \int_\theta f(y^{new}|\theta)\pi(\theta)d\theta \\
\text{Posterior Predictive Distribution:} \qquad & p(y^{new}|y^{old}) = \int_\theta f(y^{new}|\theta, y^{old})p(\theta|y^{old})d\theta
\end{aligned}
$$