1

Useful statistical and mathematical results

This is a collection of useful statistical and mathematical results.

1.1 Probability

Definition 1.1 (Mean and Variance - Discrete random variable). *Let X be a discrete random variable with support $x_1, ..., x_K$ and probability mass function (pmf) $p(x)$, then*

$$\mathbb{E}(X) = \sum_{k=1}^{K} x_k p(x_k). \tag{1.1}$$

and

$$\mathbb{V}(X) = \sum (x_i - \mathbb{E}(X))^2 p(x_i) dx. \tag{1.2}$$

Definition 1.2 (Mean and Variance - Continuous random variable). *Let X be a continuous random variable with probabilty density function (pdf) $p(x)$, then*

$$\mathbb{E}(X) = \int x p(x) dx. \tag{1.3}$$

and

$$\mathbb{V}(X) = \int (x - \mathbb{E}(X))^2 p(x) dx. \tag{1.4}$$

Lemma 1.3 (The law of iterated expectations). *Let X and Y be two random variables. Let $\mathbb{E}_X$ denote the expectation with respect to the marginal distribution for X and $\mathbb{E}_{Y|X}$ the expectation with respect to the conditional distribution of Y given X. Then*

$$\mathbb{E}_Y(Y) = \mathbb{E}_X \mathbb{E}_{Y|X}(Y|X). \tag{1.5}$$

Lemma 1.4 (Law of total variance). *Let X and Y be two random variables. Then*

$$\mathbb{V}_Y(Y) = \mathbb{E}_X \mathbb{V}_{Y|X}(Y|X) + \mathbb{V}_X \mathbb{E}_{Y|X}(Y|X). \tag{1.6}$$

Definition 1.5 (Marginal distribution - two variables). *Let X and Y be a two random variables with joint distribution $p(x, y)$.*
The marginal distribution of x when both variables are discrete is

$$p(x) = \sum_y p(x, y), \tag{1.7}$$

where the sum runs over all the values in the support of Y.
When both variables are continuous

$$p(x) = \int p(x, y) dy. \tag{1.8}$$

Definition 1.6 (Marginal distribution - n variables). *Let $X_1, \ldots, X_n$ be a set of n random variables with joint distribution $p(x_1, \ldots, x_n)$. The marginal distribution of X_i when all variables are discrete is*

$$p(x_i) = \sum_{x_1} \sum_{x_{i-1}} \cdots \sum_{x_{i+1}} \cdots \sum_{x_n} p(x_1, \ldots, x_n), \tag{1.9}$$

where the sums runs over all the values in each variable's support.
The marginal distribution of X_i when all variables are continuous is

$$p(x_i) = \int \cdots \int p(x_1, \ldots, x_n) dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n. \tag{1.10}$$

Definition 1.7 (Conditional distribution). *Let X and Y be two random variables with joint distribution $p(x, y)$ and marginal distributions $p_X(x)$ and $p_Y(y)$. Then the conditional distribution of X given Y is*

$$p(x|y) = \frac{p(x, y)}{p_Y(y)}. \tag{1.11}$$

Lemma 1.8 (Product rule - two variables). *Let X and Y be two random variables with joint distribition $p(x, y)$. Then*

$$p(x, y) = p(y)p(x|y) = p(y|x)p(x). \tag{1.12}$$

Lemma 1.9 (Product rule - n variables). *Let $X_1, \ldots, X_n$ be a set of n random variables with joint distribution $p(x_1, \ldots, x_n)$. Then*

$$p(x_1, \ldots, x_n) = p(x_1)p(x_2|x_1) \cdots p(x_n|x_1, \ldots x_{n-1}). \tag{1.13}$$

Lemma 1.10 (Law of total probability for events). *Let A and B be two events, and let B^c be the complement to B. Then*

$$p(A) = p(A|B)p(B) + p(A|B^c)p(B^c). \tag{1.14}$$

Lemma 1.11 (Bayes theorem for events). *Let A and B be two events. Then*

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(B|A)p(A)}{p(A|B)p(B) + p(A|B^c)p(B^c)}. \tag{1.15}$$

Lemma 1.12 (Bayes theorem for continuous variables). *Let X denote the data and θ a continuous parameter. Then*

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)d\theta}. \tag{1.16}$$

1.2 Standard mathematical functions

Definition 1.13 (The Gamma function).

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt \tag{1.17}$$

$$\Gamma(x) = x\Gamma(x-1) \tag{1.18}$$

$$\Gamma(n) = (n-1)! \textit{ if } n \textit{ is an integer} \tag{1.19}$$

$$\Gamma(1/2) = \sqrt{\pi} \tag{1.20}$$

Definition 1.14 (The Beta function).

$$\mathrm{B}(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt \tag{1.21}$$

$$\mathrm{B}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{1.22}$$

$$\tag{1.23}$$