Appendix C

Computation with Variances

C.1 Disclaimer

The summary shown below is based on the article [contributors, 2024]. At this point, we restrict ourselves to the category of discrete random variables.

C.2 Additional Definitions and Concepts

In order to

C.2.1 Expected Value

The expected value $E[X]$ of a given discrete random variable X and a function $g()$ is defined as

$$E[g(X)] = \sum_{x_i \in \mathcal{X}} g(x_i) * Pr(X = x_i)$$

The above definition is mostly given with $g()$ being the identity function. This then leads to

$$E[X] = \sum_{x_i \in \mathcal{X}} x_i * Pr(X = x_i)$$

The above definition can also be extended to more than one variable. Hence for random variables X and Y with a joint probability distribution $Pr(X,Y)$ and a function $h()$, we can define

$$
\begin{aligned}
E[h(X,Y)] &= \sum_{(x_i,y_i)\in\mathcal{X}\times\mathcal{Y}} h(x_i,y_i) * Pr(X=x_i, Y=y_i) \\
&= \sum_{x_i\in\mathcal{X}} \sum_{y_i\in\mathcal{Y}} h(x_i,y_i) * Pr(X=x_i, Y=y_i) \qquad \text{(C.1)}
\end{aligned}
$$

C.2.2 Properties of Expected Values

A constant factor a multiplied to X leads to

$$
E[aX] = \sum_{x_i\in\mathcal{X}} a * x_i * Pr(X=x_i) = a * \sum_{x_i\in\mathcal{X}} x_i * Pr(X=x_i) = a * E[X]
$$

The expected value of two random variables X and Y with

$$
E[X] = \sum_{x_i\in\mathcal{X}} x_i * Pr(X=x_i)
$$

and

$$
E[Y] = \sum_{y_i\in\mathcal{Y}} y_i * Pr(Y=y_i)
$$

Using the above shown random variables and assuming an existing joint probability distribution $P(X,Y)$, the expected value $E[X+Y]$ of the sum of the two random variables is given by

$$
\begin{aligned}
E[X+Y] &= \sum_{(x_i,y_i)\in\mathcal{X}\times\mathcal{Y}} (x_i+y_i) * Pr(X=x_i, Y=y_i) \\
&= \sum_{x_i\in\mathcal{X}} \sum_{y_i\in\mathcal{Y}} (x_i) * Pr(X=x_i, Y=y_i) + \sum_{x_i\in\mathcal{X}} \sum_{y_i\in\mathcal{Y}} y_i * Pr(X=x_i, Y=y_i) \\
&= \sum_{x_i\in\mathcal{X}} x_i * Pr(X=x_i) + \sum_{y_i\in\mathcal{Y}} y_i * Pr(Y=y_i) \\
&= E[X] + E[Y] \qquad \text{(C.2)}
\end{aligned}
$$

C.2.3 Covariance

The covariance ($Cov(X,Y)$) between two random variables X and Y is defined as

$$\begin{aligned}
Cov(X,Y) &= \sum_{x_i \in \mathcal{X}} \sum_{y_i \in \mathcal{Y}} (x_i - E[X])(y_i - E[Y]) * Pr(X = x_i, Y = y_i) \\
&= \sum_{x_i \in \mathcal{X}} \sum_{y_i \in \mathcal{Y}} (x_i y_i - x_i E[y] - y_i E[X] + E[X]E[Y]) * Pr(X = x_i, Y = y_i) \\
&= \sum_{x_i \in \mathcal{X}} \sum_{y_i \in \mathcal{Y}} x_i y_i * Pr(X = x_i, Y = y_i) \\
&- \sum_{x_i \in \mathcal{X}} \sum_{y_i \in \mathcal{Y}} x_i E[Y] * Pr(X = x_i, Y = y_i) \\
&- \sum_{x_i \in \mathcal{X}} \sum_{y_i \in \mathcal{Y}} y_i E[X] * Pr(X = x_i, Y = y_i) \\
&+ \sum_{x_i \in \mathcal{X}} \sum_{y_i \in \mathcal{Y}} E[X]E[Y] * Pr(X = x_i, Y = y_i) \\
&= E[XY] - E[X]E[Y] - E[Y]E[X] + E[X]E[Y] \\
&= E[XY] - E[X]E[Y] \qquad \text{(C.3)}
\end{aligned}$$

An alternative definition can be given in terms of expected values as

$$\begin{aligned}
Cov(X,Y) &= E[(X - E[X])(Y - E[Y])] \\
&= E[XY - XE[Y] - YE[X] + E[X]E[Y]] \\
&= E[XY] - 2E[X]E[Y] + E[X]E[Y] \\
&= E[XY] - E[X]E[Y] \qquad \text{(C.4)}
\end{aligned}$$

C.3 Variance

The variance ($Var(X)$) of a given discrete random variable X is defined as

$$Var(X) = \sum_{x_i \in \mathcal{X}} (x_i - E[X])^2 * Pr(X = x_i)$$

Combining both definition leads to the following alternative formulation of the variance

$$\begin{aligned} Var(X) &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + (E[X])^2] \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 \\ &= E[X^2] - 2(E[X])^2 + (E[X])^2 \\ &= E[X^2] - (E[X])^2 \end{aligned} \tag{C.5}$$

C.3.1 Variance of a Factor Times a Random Variable

The variance of a constant factor a times a random variable X is

$$\begin{aligned} Var(aX) &= E[(aX - E[aX])^2] \\ &= E[(aX - aE[X])^2] \\ &= E[a^2 * (X - E[X])^2] \\ &= a^2 * E[(X - E[X])^2] \\ &= a^2 * Var(X) \end{aligned} \tag{C.6}$$

C.3.2 Variance of a Sum

Using the alternative formulation for the sum of two random variables we have

$$\begin{aligned} Var(X + Y) &= E[(X + Y)^2] - (E[X + Y])^2 \\ &= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\ &= E[X^2] + 2E[XY] + E[Y^2] - (E[X])^2 - 2E[X]E[Y] + (E[Y])^2 \\ &= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 + 2(E[XY] - E[X]E[Y]) \\ &= Var(X) + Var(Y) + 2Cov(X, Y) \end{aligned} \tag{C.7}$$

C.4 Vector Valued Random Variables

So far, random variables such as X and Y were scalar valued. That means an instance x_i or y_i of the random variables are scalar values.

Vector-valued random variables such as $\overrightarrow{X}$ of length N can be thought of as an extension of scalar-valued random variables. An instance of $\overrightarrow{X}$ which is denoted as $\vec{x}_i$ is a vector with N elements. Hence

$$\vec{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \ldots \\ x_{iN} \end{bmatrix}$$

C.4.1 Expected Value

The expected value $E[\overrightarrow{X}]$ of the random variable $\overrightarrow{X}$ is a vector of expected values of the single elements of $\overrightarrow{X}$

$$E[\overrightarrow{X}] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \ldots \\ E[X_N] \end{bmatrix}$$

C.4.2 Variance

The variance of a vector-valued random variable is a variance-covariance matrix.

$$Var[\overrightarrow{X}] = \begin{bmatrix} Var[X_1] & Cov[X_1, X_2] & \ldots & Cov[X_1, X_N] \\ Cov[X_2, X_1] & Var[X_2] & \ldots & Cov[X_2, X_N] \\ \ldots & \ldots & \ldots & \ldots \\ Cov[X_N, X_1] & \ldots & \ldots & Var[X_N] \end{bmatrix}$$

C.5 Continuous Random Variables

All the material presented so far is valid for discrete random variables. But all relationships can also be adpated to continuous random variables. The only adaptations that we have to show here is the definition of the expected value. Hence for a continuous random variable X and a function $g()$, the expected value is defined as

$$E[g(X)] = \int_{x \in \mathcal{X}} g(x) f(x) dx$$

where $f(x)$ is the density function of the random variable X.

Compared to the discrete case, the summation was replaced by the integral and the probability function was replaced by the density. But apart from that all the above relations can also be used for continuous random variables.