Probabilistic models

Let A and B be some event in the sample space Ψ.

- $P(\Psi) = 1$
- $P(A) \geq 0$ - probability is nonnegative
- $P(A \cup B) = P(A) + P(B)$ if A and B are mutually exclusive

Conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(AB)}{P(B)} = P(B|A)P(A)$$

Bayes' theorem

$$ P(B|A) = \frac{P(A|B)P(B)}{P(A)} $$

Independence

A and B are independent if:

$$ P(A|B) = P(A) $$

$$ P(AB) = P(A)P(B) $$

Which means that knowing B does not give any information about A.

Probability density function (PDF)

$$ f_X(x) dx = P(x \geq X < x + dx) $$

The integral of the PDF must equal 1:

$$ \int_{-\infty}^{\infty} f_X(x) dx = 1 $$

Joint distribution

Consider two random variables X and Y.

The conditional PDF is the distribution of one RV when the other is fixed.

$$ f_{X|Y}(x|Y=y_0) = \frac{f_{X,Y}(x,y_0)}{f_Y(y_0)} $$

The above expression is the distribution of X given that $Y=y_0$.

Marginal PDF

$$ f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy $$

Independence

Random variables X and Y are independent if:

$$ f_{X,Y}(x,y) = f_X(x)f_Y(y) $$

Summary of information

To describe the mean of the joint distribution, we need the means of each random variable in the distribution.

To describe the spread of the joint distribution, we need the marginal variance of each RV as well as the covariance.

Matrix form

We can write two separate random variables X_1 and X_2 in a matrix:

$$ \mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} $$

The expectation is given by:

$$ E[\mathbf{X}] = \begin{bmatrix} \mu_{X_1} \\ \mu_{X_2} \end{bmatrix} = \mathbf{\mu_X} $$

$$ \tilde{\mathbf{X}} = \begin{bmatrix} X_1 - \mu_{X_1} \\ X_2 - \mu_{X_2} \end{bmatrix} = \mathbf{X} - \mathbf{\mu_X} $$

Covariance matrix

$$ \mathbf{C_{XX}} = E\left[\tilde{\mathbf{X}}\tilde{\mathbf{X}}^T\right] = \begin{bmatrix} \sigma_{X_1X_1} & \sigma_{X_1X_2} \\ \sigma_{X_1X_1} & \sigma_{X_1X_2} \end{bmatrix} $$

Bivariate Gaussian: matrix generalization of the Gaussian

$$ f_\mathbf{X}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^2 \mathrm{det} \mathbf{C_{XX}}}} \mathrm{exp} \left[-\frac{1}{2} (\mathbf{X} - \mathbf{\mu_X})^T \mathbf{C_{XX}}^{-1} (\mathbf{X} - \mathbf{\mu_X}) \right] $$

Correlation coefficient

$$ \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} $$

Effect of coordinate transformation

Consider two random variables X_1 and X_2.

We can transform the coordinates from x_1, x_2 to a different set of coordinates z_1, z_2 by multiplying by a transformation matrix $\mathbf{M}$.

$$ \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \mathbf{M} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} $$

The mean/expectation stays in the same location, just with the transformation applied to it, because expectation is a linear operator.

$$ \begin{bmatrix} \mu_{Z_1} \\ \mu_{Z_2} \end{bmatrix} = \mathbf{M} \begin{bmatrix} \mu_{X_1} \\ \mu_{X_2} \end{bmatrix} $$

Use the following relation to find the new covariance matrix:

$$ \tilde{\mathbf{Z}} = \mathbf{M} \tilde{\mathbf{X}} $$

The new covariance matrix $\mathbf{C_{ZZ}}$ is:

$$ \mathbf{C_{ZZ}} = \mathbf{M}\mathbf{C_{XX}}\mathbf{M}^T $$

Effect of shifting and scaling

Let $V=\alpha (X-\beta)$ and $W = \gamma(Y-\delta)$.

For these new variables:

$$ \mu_V = \alpha (\mu_X - \beta) $$ $$ \sigma_V^2 = \alpha^2 \sigma_X^2 $$ $$ \sigma_{VW} = \alpha \gamma \sigma_{XY} $$

The aforementioned correlation coefficient is invariant to shifting and scaling.