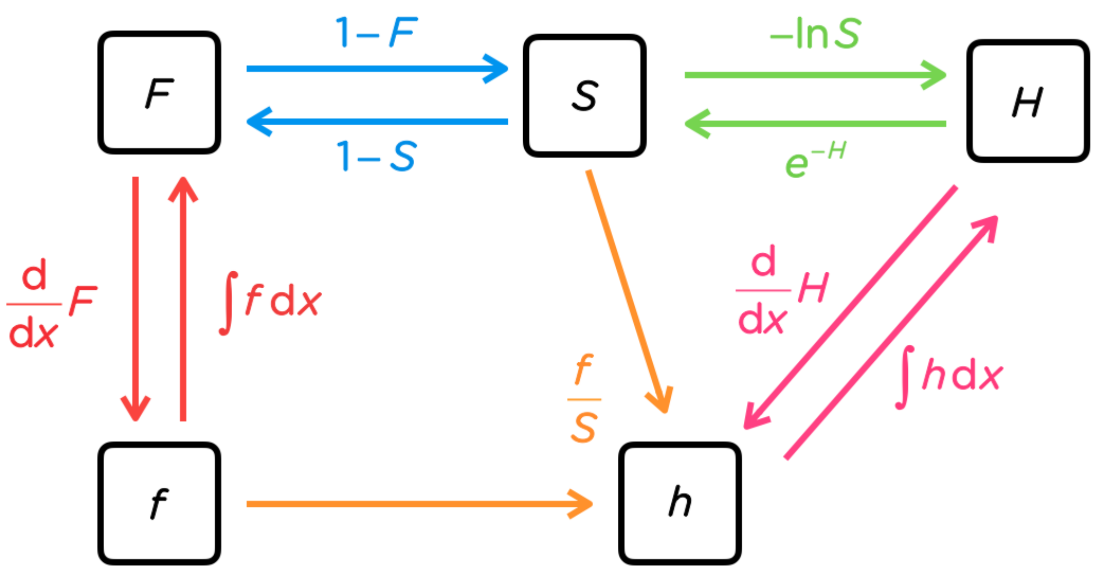
Summary 10M

Functions

This chart summarizes the relationships between the various functions of a continuous random variable.



Percentiles

The $100p^{\text{th}}$ percentile of a random variable X is any value π_p that satisfies these two properties:

- $\Pr(X < \pi_p) \le p$
- $\Pr(X \le \pi_p) \ge p$

If X is continuous, the $100p^{\text{th}}$ percentile is the value π_p such that $F(\pi_p) = p$.

Mode

The mode of a random variable is the value that produces the largest PMF or PDF.

If the random variable X is continuous, then

$$f'(x) = 0 \quad \Rightarrow \quad x = \text{mode}$$

Moments

- The k^{th} raw moment of X is defined as

$$\mu'_k = \mathrm{E}\left[X^k\right]$$

- The 1^{st} raw moment is the mean, which is usually denoted by μ.

$$\begin{aligned} \mu &= \mathrm{E}[X] \\ &= \sum_{\text{all } x} x \cdot p(x) && \text{(discrete)} \\ &= \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x && \text{(continuous)} \end{aligned}$$

- The k^{th} central moment of X is defined as

$$\mu_k = \mathrm{E}\left[(X-\mu)^k\right]$$

- The 2^{nd} central moment is the variance, which is usually denoted by $\mathrm{Var}[X]$ or σ^2.

$$\begin{aligned} \sigma^2 &= \mathrm{Var}[X] \\ &= \mathrm{E}\left[(X-\mu)^2\right] \\ &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \end{aligned}$$

- The standard deviation, denoted σ, is the square root of the variance.

$$\sigma = \sqrt{\text{Var}[X]}$$

- The coefficient of variation is the ratio of the standard deviation to the mean.

$$CV = \frac{\sigma}{\mu}$$

- The skewness of a distribution is:

$$\text{Skewness} = \frac{\mu_3}{\sigma^3}$$

- The kurtosis of a distribution is:

$$\text{Kurtosis} = \frac{\mu_4}{\sigma^4}$$

Conditional Distribution

- Bayes' Theorem:

$$\Pr(A \,|\, B) = \frac{\Pr(B \,|\, A) \Pr(A)}{\Pr(B)}$$

- Based on the Law of Total Probability,

$$\Pr(X = x) = \text{E}_Y[\Pr(X = x \mid Y)]$$

- Based on the Law of Total Expectation,

$$\mathrm{E}_X[X] = \mathrm{E}_Y[\mathrm{E}_X[X \mid Y]]$$

- Based on the Law of Total Variance,

$$\mathrm{Var}_X[X] = \mathrm{E}_Y[\mathrm{Var}_X[X \mid Y]] + \mathrm{Var}_Y[\mathrm{E}_X[X \mid Y]]$$

Independence

Events A and B are independent if and only if the following equation is true:

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$$

If random variables X and Y are independent, then

$$\Pr(X = x,\, Y = y) = \Pr(X = x) \cdot \Pr(Y = y)$$

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

$$\mathrm{E}[g(X) \cdot h(Y)] = \mathrm{E}[g(X)] \cdot \mathrm{E}[h(Y)]$$

Empirical Distributions

The empirical distribution is a discrete distribution based on a sample of size n that assigns probability $\frac{1}{n}$ to each data point.

- Empirical distribution function

$$F_n(x) = \frac{\text{Number of observations} \le x}{n}$$

- Empirical $100p^{\text{th}}$ percentile

$$\pi_p = x_{(\lceil np \rceil)}$$

- Sample mean

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- Biased sample variance

$$\begin{aligned} \text{Var}[X] &= \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} \\ &= \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2 \end{aligned}$$

- Unbiased sample variance

$$\begin{aligned} s^2 &= \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \\ &= \frac{n}{n-1} \cdot \text{Var}[X] \end{aligned}$$