Introduction to (Bayesian) Estimation
MAE 5020

Review of Probability theory and Bayes theorem

Oklahoma State University

August 11, 2023

Basic probability theory and random variables[1]

- Probability theory in "problem solving":
  The outcome of any one occurrence *trial* of the phenomenon (*experiment*) cannot be predicted, but is rather governed by randomness.
- Example: determine the number of times an even number will come up in a thousand rolls of a fair die.
  1. prior probabilities:
  2. operations:
  3. prediction:

  Probability theory is about the 2nd step. Step 1 and 3 belong to statistics.

[1] Stochastic Processes and Filtering theory by Jazwinski

Assign Prior Probabilities

- Relative frequency from experiments:
- Classical method:
- Measure of belief:

Fundamental concepts in Probability

- *probability space* Ω, *elements* ω (samples or experiment outcomes). Certain subsets of Ω (collection of outcomes) are called *events*.
- We assign probabilities to events via a *probability function* $Pr\{\cdot\}$.
- Example: rolling of a die

Probability axioms

1. For any event Λ, $Pr\{\Lambda\} \geq 0$.
2. $Pr\{\Omega\} = 1$.
3. Any countable sequency of *mutually exclusive* events Λ_1, Λ_2, $\cdots$ satisfy

$$Pr\{\cup_{i=1}^{\infty}\Lambda_i\} = \sum_{i=1}^{\infty} Pr\{\Lambda_i\}.$$

Implication:

- $Pr\{\emptyset\} = 0$
- $Pr\{\Lambda\} = 1 - Pr\{\Omega - \Lambda\} \leq 1$
- $Pr\{\Lambda_1 \cup \Lambda_2\} \leq Pr\{\Lambda_1\} + Pr\{\Lambda_2\}$
- If $\Lambda_1 \subset \Lambda_2$, $Pr\{\Lambda_1\} \leq Pr\{\Lambda_2\}$.

Random variables (rv)

- Scalar rv: a scalar variable x may take different values each time it is sampled/measured. If these values bear no deterministic (fixed) relationship to the sampling process, x is a scalar rv.
- For now, x only takes discrete values, i.e., discrete scalar rv.
- Example: scoop of jelly beans in a jar.

Statistics of x

- Determined by experiments.
- Given N experiments, the frequency of occurrence of a specific x provides an estimate of the probability that x will occur in a future event:
- $Pr(x_i)$:

x takes continuous values

- Example: dispensing juice. x represents the volume of each glass.

- Probability density function:

Statistics: expectation and its property

Statistics: nth moment

Example: uniform distribution

Consider a rv uniformly distributed in the interval $[0, T]$. Calcuate its mean and variance.

Example: Gaussian (normally distributed) rv

- A rv is Gaussian if its pdf is given by $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-m)^2}{\sigma^2}}$.
- Mean, Variance:

Two random variables

- The probabilities of each rv takes a certain value may be independent or related to the other rv.
- Example: jelly beans have two colors (orange and black)

Joing probability $Pr(x_i, y_i)$

Probability of x_i orange jelly beans and y_i black jelly beans are scooped every time.

- Whehter or not x_i and y_i are related,

Conditional probabilities

Bayes theorem

Example

$Pr(rain) = p_r$, $Pr(cloudy|rain) = p_c$, $p(cloudy|not\ rain) = p_n$, what is $p(rain|cloudy)$?

$$
\begin{aligned}
p(rain|cloudy) &= \frac{p(cloudy|rain)p(rain)}{p(cloudy)} \\
&= \frac{p_c p_r}{p(cloudy|rain)p(rain) + p(cloudy|not\ rain)p(notrain)} \\
&= \frac{p_c p_r}{p_c p_r + p_n(1 - p_r)}
\end{aligned}
$$

$p_c + p_n = ?$

Continuous variables

Vector rv

Vector rv

Independence vs. uncorrelated

Example

Consider two random variables y_1 and y_2 defined by $y_1 = \sin 2\pi x$ and $y_2 = \cos 2\pi x$, where x is uniformly distributed on $[0, 1]$. Show that $\mathbb{E}(y_1) = \mathbb{E}(y_2) = \mathbb{E}(y_1 y_2) = 0$.

Show that y_1, y_2 are uncorrelated. Are they independent?

n jointly normally distributed rv $x_1, \cdots, x_n$

$$p(x) = \frac{1}{\sqrt{(2\pi)^n |P_x|}} e^{-\frac{1}{2}(x - \mathbb{E}(x))^T P_x^{-1} (x - \mathbb{E}(x))}$$