COSC 3P99 Notes

Tyler McDonald

January 2023

1 Probability - Univariate Models

Attribution: These notes follow Kevin P. Murphy's *Probabilistic Machine Learning - An Introduction*, (2022, MIT Press). The draft PDF was accessed via probml.ai.

1.1 Introduction

Probabilistic Interpretations - We differentiate how we look at and consider probability into two distinct interpretations - the **frequentist** interpretation and the **Bayesian** interpretation:

- **Frequentist Interpretation** - If we flip a fair coin - where the probability of both cases are equal (expressed mathematically as $Pr(X) = Pr(Y) = 0.5$), we consider the case of heads to occur in about half of all present and future cases. The key point here is that we consider the **summation of all cases**, not just the result of any one case.
- **Bayesian Interpretation** - If we flip the same fair coin, we consider the **uncertainty of the current case**, and express ignorance of the future cases. To simplify this, Bayesian inference of an outcome considers each case as if it were both the final case and determinant case - all cases prior are meaningless sample sets.

We primarily utilize Bayesian interpretation for one-off events - if an event can happen exactly once, or not at all - in a sense, an atomic event - we want to quantify our **uncertainty** of either outcome, not the odds that across 10 samples it happens.

Uncertainties - We separate the idea of uncertainty into separate classifications depending on the uncertainty we are aiming to quantify:

- **Epistemic Uncertainty** - Derived from *epistemology*, or the study of knowledge, and referred to colloquially as **model uncertainty**; epistemic uncertainty arises from ignorance of hidden causes or mechanisms generating our data. Think: do we consider the effects of flipping coins of different material composition in our experimentation?

- **Aleatoric Uncertainty** - Derived from *alea iacta est*, or a Latin dice game of chance, and referred to colloquially as **data uncertainty**; aleatoric uncertainty arises from intrinsic variability - the idea of true randomness dictates a level of intrinsic variability. Think; can we perfectly predict 15 fair coin flips without seeding the outcome?

1.2 Events

Definition 1. *An* ***event****, denoted by the binary variable A, is defined as some state of the world that either holds or does not hold.*

We use $Pr(A)$ as the mathematical notation for "the probability of A" iff $0 \leq Pr(A) \leq 1$. A state of $Pr(A) = 0$ denotes an impossible event, while a state of $Pr(A) = 1$ denotes a certain event.

We use $Pr(\bar{A})$ to denote the probability of an event *not* occurring. Thus, it holds that $Pr(\bar{A}) = 1 - Pr(A)$.

Definition 2. *Let A and B be two events. The* ***joint probability of A and B****, denoted $Pr(A, B)$, is equal to the intersection of the probability values of each event, denoted $Pr(A \wedge B)$.*

If A and B are independent events - so to speak, B does not rely on an input value equal to the output of A - then this joint probability is equal to the product of the probability of A and B, or $Pr(A)Pr(B)$.