STAT 305 2024S2 **Instructor: Vivian Meng**
Lecture 6.1 **Date: July 16, 2024**

Learning Goals

G1. Bayesian vs Frequentist interpretation of probability
G2. Bayesian paradigm

Bayesian vs Frequentist interpretation of probability

So far, we have primarily interpreted "probability" from a **frequentist** perspective; that is, given a random experiment with a sample space Ω, and an event $E \subseteq \Omega$, then $P(E)$ represents the long run proportions of the times event E happens if we could repeat the random experiment unlimited number of times.

In Chapter 6 of the course notes we venture into **Bayesian** inference. Here, probability has a subjective interpretation. Under the Bayesian paradigm, probability represents one's **subjective assessment of uncertainty** regarding some event happening.

It should be easy to agree that:

- $P(E) = 1$ means
- $P(E) = 0$ means
- Let X represent the class average for Midterm 1 in STAT 305 for the 2024S2 term. $P(X \in [75, 80]) = ?$

  Clearly, $\{X \in [75, 80]\}$ is an observable event that may or may not happen.

For subjective prob to represent one's uncertainty, it makes sense that the more uncertain we feel towards an event happening, the smaller the probability associated with that event. So, relatively speaking, if E_1 and E_2 are two events such that we feel more certain of E_2 occuring than E_1, then $P(E_2) > P(E_1)$.

Operationalization of subjective probability

As to how to **coherently** assign a value to $P(E)$ for events that is less than certain to happen/not happen, the Italian probabilist Bruno De Finetti proposes the following way for operationalization of subjective probability.

You are to pretend to be the "house" who sets the odds table (as below) for betting on or against some event. Furthermore, you do not know if the gambler will choose to bet on or against the event.

Event	Offered odds	Betting **on** the Event	Betting **against** the event
E	pays a to b	gains a dollars for every b dollars wagered if E happens, otherwise lose b dollars	gains b dollars for a dollars wagered if event does **not** happen, otherwise lose a dollars

Once you have specified the betting odds for an event, then your subjective probability is $P(E) = b/(a + b)$.

The following articles explains some ways to interpret subjective probability.

- `https://en.wikipedia.org/wiki/Coherence_(philosophical_gambling_strategy)`
- `https://en.wikipedia.org/wiki/Dutch_book`
- (Chapter 2 of) Bernardo, J.M. and Smith, A.F.M. (1994) Bayesian Theory. John Wiley and Sons, Chichester.

Bayesian paradigm

In terms of inferring a model parameter, as we are *uncertain* about the true value of model parameter we describe our uncertainty via a probability distribution.

> In STAT 305, we view parameter value π as realization of a RV "Π" (upper case pi), where probability distribution of Π describes subjective uncertainty regarding what the parameter value might be.
>
> Other common greek letters for parameters:

Beyond STAT305 unfortunately you will probably see the same lower case greek letters representing both the parameter as a RV, and one of its realization/value.
The Bayesian principle says that upon observing data our uncertainty regarding the parameter of interest should update to a **conditional distribution** for the parameter given observed data, calculated following the general structure of Bayes rule.

> Let $f_{X|\Pi}(x|\pi)$ be the pmf of observable RV X conditional on the value of model parameter being π. If we believe the possible values of the model parameter is given by a countable set $\mathcal{R}_\Pi$, then the pmf $p_\Pi(\pi)$ represents one's subjective "uncertainty" regarding the value of the parameter *prior* to seeing the data. Then, Bayes rule says
>
> $$p_{\Pi|X}(\pi|X = x) = \frac{f_{X|\Pi}(x|\pi)p_\Pi(\pi)}{\sum_{\mathcal{R}_\Pi} f_{X|\Pi}(x|\pi)p_\Pi(\pi)}$$
>
> which gives the pmf of the *posterior* distribution of Π conditioning on observing $X = x$. (*We will relax the assumption of countable $\mathcal{R}_\Pi$ in the next lecture)

Lecture 6.1 Student Activities

1. Operationalization of subjective probability.

   (a) Let T represent the maximum temperature that occurs in Vancouver on August 1st, 2024, measured in $^\circ C$. You are the "house", and must complete the following betting odds table such that you are indifferent to a gambler choosing to bet on or against any of the events listed according to the offered odds.

Event	Offered odds
$T \leq$	pays 9 to 1
$T \leq$	pays 2 to 1
$T \leq$	pays 1 to 1
$T \leq$	pays 1 to 2
$T \leq$	pays 1 to 9

   (b) Based on the above odds table, sketch a CDF which represents your **subjective** probability associated with T.

2. Let $f_{X|\Lambda}(x|\lambda)$ be the pmf of observable RV X conditional on model parameter taking value λ, and assume possible values of the model parameter is give by a countable set $\mathcal{R}_\Lambda$, and that pmf $p_\Lambda(\lambda)$ characterizes our prior distribution for Λ (i.e. uncertainty regarding model parameter before collecting data).

   (a) Write down the joint probability of $\{X = x \text{ and } \Lambda = \lambda\}$

   (b) Write down an expression for the marginal probability of $X = x$

   (c) Interpret the marginal distribution $P(X = x)$ in terms of model "averaging".

   (d) Under the Bayesian paradigm, $P(X = x)$ is a subjective probability. Interpreting this in terms of betting odds for an ideal rational/coherent person.

