603 WEEK1 NOTES

Probability Theory in Forecasting

- **Importance of Probability in Forecasting**: Essential to establish a confidence bound.
  - Example: Investing $100 in stock, predicting future worth with certain probabilities.
  - Need to understand the underlying statistical distribution to make accurate predictions.

1. Confidence Bound Statements:

- Example of confidence in value after a period (e.g., "95% chance value will be between $300 and $500 in 10 years").
- Key points:
  - Forecasting requires probability of forecast being correct.
  - Confidence intervals are essential.

Individual and Compound Probabilities

- **Probability of Individual Events**:
  - E.g., Probability of a single coin flip yielding heads.
- **Compound Probabilities**:
  - E.g., Probability of two coin flips yielding heads both times.
- **Application**:
  - Probability in financial models to estimate events with multiple variables (e.g., economic recession impact on bonds).

Measuring Probability with Random Variables

- **Random Variables**:
  - A random variable maps the set of outcomes to real numbers.
  - Types:
    - Discrete: Countable outcomes (e.g., dice roll).
    - Continuous: Any value within a range (e.g., stock return percentages).

Discrete and Continuous Random Variables

- **Discrete Random Variables**:
  - Probability for specific values $P(X = x_i) = p_i$.
- **Continuous Random Variables**:
  - Probability Density Function (PDF) for ranges, $P(r_1 < X < r_2) = p$.

Probability Density Functions (PDF)

- For a continuous variable, the PDF describes the likelihood of outcomes within intervals.
  - Example integral for probability between two points r_1 and r_2:

$$P(r_1 < X < r_2) = \int_{r_1}^{r_2} f(x)dx$$

Cumulative Distribution Functions (CDF)

- CDF gives probability that a variable is less than a certain value.
  - Example:

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)dx$$

Inverse Cumulative Distribution Functions

- Inverse CDF (quantile function): Determines value given a probability.
  - 95th percentile is $F^{-1}(0.95)$.

Normal Distribution CDF

- **Normal Distribution Characteristics**:
  - Parameters include mean μ and standard deviation σ.
  - Standard Normal CDF graph with different means and variances.

Sample Problems

####1. **CDF Example**: - Given $F(a) = \frac{a^2}{100}$, calculate values to meet specific conditions.

####2. **Exponential Distribution**: - CDF $F(x; \lambda) = 1 - e^{-\lambda x}$ for $x \geq 0$.

Probability of Mutually Exclusive Events

- **Formula**:
  - $P(A \cup B) = P(A) + P(B)$ for exclusive events.
  - Example calculation for stock returns.

Independent Events and Joint Probability

- **Joint Probability of Independent Events**:
  - For independent events W and M, $P(W \text{ and } M) = P(W) \cdot P(M)$.

Probability Matrices

- Useful for showing joint probabilities of two variables.
  - Example: Stock grading matrix for outperforming/underperforming bonds.

Conditional Probability

- **Conditional Probability Formula**:
  - $P(M|W) = \frac{P(M\cap W)}{P(W)}$.
  - Example calculation given specific outcomes for stock upgrades.

Independence

- **Definition of Independence**:
  - If $P(M|W) = P(M)$, then M and W are independent.

Applications of Bayesian Analysis

Bayesian Analysis 应用

Overview: Bayesian analysis enables forecasting in situations with limited data.

概述：贝叶斯分析可以在数据不足的情况下进行预测。

- **Key Points**:
  - Bayesian analysis allows decision-making under uncertainty.
  - In low-data scenarios, Bayesian priors can compensate by integrating prior knowledge.
- **要点**：
  - 贝叶斯分析支持在不确定性下的决策。
  - 在数据稀缺的情境中，贝叶斯先验通过整合先验知识来弥补不足。

Bayes' Theorem

贝叶斯定理

Bayes' theorem is a foundational principle in Bayesian analysis, allowing us to update probabilities based on new evidence.

贝叶斯定理是贝叶斯分析的基础原理，允许我们根据新证据更新概率。

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

- **Explanation**:
  - $P(A|B)$: Posterior probability (Updated probability of A given B)
  - $P(B|A)$: Likelihood (Probability of B given A)
  - $P(A)$: Prior probability (Initial belief about A)
  - $P(B)$: Marginal probability (Normalization factor)
- **解释**：
  - $P(A|B)$: 后验概率（给定 B 后的 A 的更新概率）

- $P(B|A)$: 似然（给定 A 后 B 的概率）
- $P(A)$: 先验概率（关于 A 的初始信念）
- $P(B)$: 边际概率（归一化因子）

Illustration: Bond Default

示例：债券违约

Suppose we want to calculate the probability of a bond default given certain economic trends. Defaults may be positively correlated due to shared economic factors.

假设我们要计算在特定经济趋势下债券违约的概率。由于共享的经济因素，违约可能呈正相关。

- **Assumptions**:
  - Probability of Bond A or Bond B defaulting individually is 10%.
  - Joint probability for both bonds defaulting is 6%.
- **假设**:
  - 债券 A 或 B 单独违约的概率为 10%。
  - 两者共同违约的联合概率为 6%。

Joint and Conditional Probabilities

联合概率和条件概率

Joint probability $P(A \cap B)$ and conditional probability $P(A|B)$ help in understanding dependencies.

联合概率 $P(A \cap B)$ 和条件概率 $P(A|B)$ 有助于理解变量间的依赖关系。

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Deriving Bayes' Theorem from Conditional Probability

从条件概率导出贝叶斯定理

Bayes' theorem can be derived by rearranging conditional probabilities.

通过重组条件概率可以推导出贝叶斯定理。

1. **Starting from Joint Probability**:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

2. **Rearrange**:

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

Prior and Posterior Beliefs

先验和后验信念

In Bayesian inference, we update prior beliefs (initial assumptions) to posterior beliefs using Bayes' theorem.

在贝叶斯推理中，我们通过贝叶斯定理将先验信念（初始假设）更新为后验信念。

- **Prior** $P(A)$: Initial assumption before observing data.
- **Posterior** $P(A|B)$: Updated belief after observing evidence B.
- **先验** $P(A)$：在观察数据前的初始假设。
- **后验** $P(A|B)$：在观察证据 B 后的更新信念。

Sample Problem: Fund Managers

示例问题：基金经理

Calculate the probability of a manager being skillful given that they passed a certain performance test. 计算某经理通过特定绩效测试后，具备专业能力的概率。

- Given:
  - Skillful managers are 1 in 100.
  - Test accuracy is 99% for skillful and 97% for unskillful.

$$P(\text{Skillful}|\text{Positive Test}) = \frac{P(\text{Positive Test}|\text{Skillful}) \cdot P(\text{Skillful})}{P(\text{Positive Test})}$$

Bayes versus Frequentists

贝叶斯与频率派比较

Bayesian and Frequentist interpretations differ fundamentally in how they treat probabilities and handle uncertainty.

贝叶斯派和频率派在处理概率和不确定性方面有根本性的差异。

- **Bayesian**: Probability represents a degree of belief.
- **Frequentist**: Probability is the long-run frequency of events.
- **贝叶斯派**：概率代表信念的程度。
- **频率派**：概率是事件在长期中出现的频率。

Highest Posterior Density Interval

最高后验密度区间

The Highest Posterior Density (HPD) interval represents the range containing a specified portion (e.g., 95%) of the posterior distribution.

最高后验密度（HPD）区间表示包含后验分布的指定部分（例如 95%）的范围。

- **Interpretation**:
  - HPD provides an interval estimate for parameters, capturing the most credible values.
- **解释**：
  - HPD 为参数提供区间估计，捕获最可信的值。

Summary

总结

In Bayesian analysis, we use prior beliefs and observed data to make probability-based decisions. This method is particularly useful in situations with limited data, allowing for structured uncertainty handling.

在贝叶斯分析中，我们使用先验信念和观察数据来做基于概率的决策。这种方法在数据有限的情况下特别有用，能够有条理地处理不确定性。