这篇文章的核心是基于潜在结果框架的方法，介绍了如何从观测数据中推断因果关系。

**第一部分：因果推断的背景与挑战**

**1. 因果推断的吸引力**

• 相比随机对照实验（Randomized Controlled Trials, RCTs），观测数据因果推断更具有吸引力，主要原因是：

• **成本更低**：不需要设计实验。

• **数据规模更大**：尤其是在现代数据环境下，大规模观测数据随处可得。

**2. 面临的挑战**

1. **如何从观测数据中获得反事实结果？**

• 我们只能观测到一个处理下的结果，但未处理下的结果（反事实结果）是不可见的。

2. **混杂因素和选择偏差**：

• 混杂因素是影响处理分配和结果的共同变量。

• 选择偏差来源于观测数据的非随机分配。

**第二部分：潜在结果框架**

潜在结果框架是因果推断的基础，用于刻画每个个体在不同处理条件下的潜在结果。

• **潜在结果（Potential Outcomes）**：

• 描述每个个体在接受处理和未接受处理时的结果。

• **反事实结果（Counterfactual Outcomes）**：

• 指未被观测到的潜在结果。

**三大关键假设**

1. **SUTVA（稳定单位处理值假设）**：

• 假设单位之间没有干扰，且每种处理只有一个版本。

2. **可忽略性**：

• 给定背景变量，处理分配与潜在结果独立。

• 通俗来说，处理的分配是随机的。

3. **正值性**：

• 对所有背景变量，每种处理的概率都是正的。

**第三部分：主要方法分类**

**1. 基于三大假设的方法**

这些方法依赖潜在结果框架的假设，通过调整样本分布或建模特定关系，推断因果效应。

**（1）重加权方法**

• **核心思路**：通过对样本赋予权重，调整处理组和对照组的协变量分布。

• **方法示例**：

1. **逆概率加权（IPW）**：

• 使用倾向分数 （接受处理的概率）作为权重，平衡处理组和对照组。

2. **Doubly Robust Estimator**：

• 结合倾向分数和结果模型，即使一个模型错误，仍能得到一致估计。

**（2）匹配方法**

• **核心思路**：通过找到相似的样本对，模拟随机对照实验。

• **方法示例**：

1. **最近邻匹配（NN Matching）**：

• 按最小距离匹配处理组和对照组样本。

2. **Kernel Matching**：

• 使用核函数赋权，估计反事实结果。

**（3）树模型方法**

• **核心思路**：通过树模型分区，识别异质性因果效应。

• **方法示例**：

1. **CART**：

• 将数据分区，并计算每个分区的平均处理效应。

2. **BART**：

• 在贝叶斯框架下，使用多棵树的加和模型估计因果效应，捕捉复杂非线性关系。

**（4）表征学习方法**

• **核心思路**：通过学习新的特征表示，平衡处理组和对照组的分布。

• **方法示例**：

• 平衡表征学习：通过最小化组间分布差异（如 Wasserstein 距离），同时优化预测误差。

**2. 放宽假设的方法**

这些方法在一定程度上放宽了潜在结果框架的假设，适应更复杂的因果问题。

**（1）处理未观测混杂因素**

• **工具变量法（Instrumental Variables, IV）**：

• 利用与结果无直接关联但与处理分配相关的工具变量估计因果效应。

• **前门准则（Front-door Criterion）**：

• 通过中介变量分解因果路径，绕过未观测混杂因素的干扰。

**（2）非参数估计与机器学习方法**

1. **深度学习方法**：

• **GANs**：通过生成平衡的潜在分布，减少混杂影响。

2. **强化学习**：

• 用于优化动态因果决策。

3. **随机森林与因果森林**：

• 扩展随机森林以估计个体化处理效应（ITE）。

4. **元学习**：

• 在多任务或动态环境中快速适应，提高因果推断效率。

**第四部分：总结与思考**

1. **贡献**：

• 本文系统总结了因果推断的核心方法，尤其是潜在结果框架下的实践。

• 提出了放宽假设的方法，为复杂问题提供了更多可能性。

2. **未来研究方向**：

• 如何更高效地结合机器学习与因果推断？

• 在更复杂的动态环境中如何实现高鲁棒性的因果效应估计？

**结束语**

今天分享的内容是《A Survey on Causal Inference》的核心，希望大家对因果推断的潜在结果框架和主要方法有初步了解。这些方法虽然看似复杂，但它们为解决实际问题提供了强大的工具。如果有任何问题，我们可以进一步探讨，谢谢！