7.1 统计推断 2019年8月15日09点53分

定义7.1.2 统计推断 统计推断是产生关于统计模型部分或所有概率的陈述过程。

定义7.1.3 参数/参数空间 在统计推断问题中，用于决（确）定目标随机变量联合分布的特征或组合特征被称为分布参数. 参数所有可能值的集合或参数向量被称为参数空间.

定义7.1.4 统计 假设目标可观测的随机变量是. 设r是任意n个实数变量的实值函数. 则随机变量被称为统计.

7.2 先验和后验分布

定义7.2.1 先验分布/p.f/p.d.f 假设有一个参数为的统计模型. 如果是随机的，则在观察其它目标随机变量之前为分配的分布被称为先验分布. 如果参数空间是可数的，则先验分布是离散的并且它的p.d.f.被称为先验p.f. 如果先验分布是连续的，则它的p.d.f.被称为先验p.d.f. 我们使用符号来标记先验p.f.或先验p.d.f.作为的函数.

当我们将参数视为随机变量时，名称“先验分布”仅仅是参数边际分布的另一个名称。

定义7.2.2 后弦分布/p.f./p.d.f. 考虑一个参数为的统计推断，并且已经观察到随机变量. 在给定条件下的条件分布被称为后验分布. 在给定条件下的条件p.f.或p.d.f.被称为的后验p.f.或后验p.d.f.并且被标记为.

当我们将参数视为随机变量时，名称“后验分布”仅仅是在给定数据情况下参数条件分布的另一个名称。

定理7.2.1假设某个分布的p.d.f或p.f.为, 从该分布中选取n个随机变量. 同样假设参数值未知且的先验p.d.f.或p.f.为. 则的后验p.d.f.或p.f.为

其中是的边际联合p.d.f或p.f.

定义7.2.3 似然函数 当随机样点中观察联合p.d.f.或联合p.f. 在给定值条件下作为函数，该函数被称为似然函数.

7.3 共轭先验分布 2019年8月28日09点58分

定理 7.3.1 假设从参数为的伯努利分布中选取n个随机样本, 其中. 同样假设先验分布是参数为的贝塔分布. 则在给定条件下后验分布是参数为和的贝塔分布.

定义 7.3.1 共轭族/超参数 设,是给定下的条件i.i.d，并且它们的共同p.f.或p.d.f.为. 设是参数空间上的可能分布族. 假设先验分布不论从怎样选取，无论选取多少观察样本，同样，无论观察值是什么, 后验分布始终是成员. 则被称为来自分布的样本的先验分布共轭族. 族也被称为自分布的闭采样(closed under sampling). 最终，如果中的分布被其它参数参数化, 则于先验分布关联的参数被称为先验超参数，于后验分布关联的参数被称为后验超参数.

定理7.3.2 假设从参数为的泊松分布中选取n个随机样本, 其中未知. 同样假设先验分布是参数为的伽马分布. 则在给定条件下后验分布是参数为和的伽马分布.

定理7.3.3 假设从均值, 方差的正态分布中选取n个随机样本，其中和未知. 同样假设先验分布是均值为, 方差为的正态分布. 则在给定条件下后验分布是均值为, 方差为的正态分布. 其中

定理7.3.4 假设从参数为的指数分布中选取n个随机样本, 其中未知. 同样假设先验分布是参数为的伽马分布. 则在给定条件下后验分布是参数为和的伽马分布.

定义7.3.2 不正常先验(Improper Prior) 设是一个非负函数，其定义域包含统计模型的参数空间. 假设. 如果我们假装是先验p.d.f., 则我们对于使用了一个不正常先验.

7.4 贝叶斯估算器 2019年8月29日10点18分

定义7.4.1 估算器/估值 设是一组观察数据，其联合分布是由从实数子集中取值的参数决定的. 参数的估算器是一个实值函数. 如果被观察到，则被称为的估值.

定义7.4.2 损失函数 损失函数是二元变量的实值函数, 标记为, 其中是一个实数. 该函数的意义是当参数等于且估值等于时的统计损失.

定义7.4.3 贝叶斯估算器/估值 设是一个损失函数. 对于的每一个可能值, 设是一个值使得最小化. 则被称为的贝叶斯估算器. 一旦=被观察到，被称为

贝叶斯估值.

定义7.4.4 平方误差损失函数

推论7.4.1 设是一个实值参数. 假设平方误差损失函数(7.4.4)被使用，并且后验均值是有限的. 则贝叶斯估算器是.

定义7.4.5 绝对误差损失函数

推论7.4.2 当绝对误差损失函数被使用时，实值参数的贝叶斯估算器等于后验分布的中位数.

定义7.4.6 一致估算器 在情况下，一组估算器概率收敛于被估算参数的未知值被称为一致估算序列。

定义7.4.7 估算器/估值 设是一组观察数据，其联合分布是由从维空间子集中取值的参数决定的. 设是从到维空间的函数. 定义. 估算器是一个从维空间取值的函数. 如果被观察到，则被称为的估值.

7.5 最大似然估算器 2019年8月30日10点50分

定义7.5.2 最大似然估算器/估算 对每一个可能的观察向量, 设在情况下使得似然函数最大化, 设为定义在这种方式下的的估算器. 估算器被称为的最大似然估算器. 当被观察时，值被称为的最大似然估值.

7.6 最大似然估算器的属性 2019年9月2日10点43分

定理7.6.1 M.L.E.的不变性 如果是的最大似然估算器并且g是一对一函数，则是的最大似然估算器.

定义7.6.1 M.L.E.函数 设是参数的任意函数, 并且设G是在函数g下的图像. 对于每一个, 定义和

最终，定义的M.L.E，其中

定理7.6.2 设是的M.L.E. 设是关于的函数. 则的M.L,E是.

定义7.6.3 矩方法 假设从k维参数的分布中选取n个随机变量, 并且该分布至少包含k个有限矩. 当 设. 假设函数是一对一函数. 设是其反函数，对于所有的，

通过定义样本矩, 其中 矩估算器方法是.

定理7.6.3 假设是i.i,d, 其分布是由k维参数向量决定的. 假设该分布对于所有前k个矩存在且均为有限值. 同样假设定义在7.6.3中的反函数M是连续的. 则基于矩估算器方法的序列是估算器是连贯序列.

7.7 充分统计量 2019年9月4日10点21分

定义7.7.1 充分统计量 设是随机样本，来自参数为的分布. 设T是一个统计量. 假设对于每一个和T的每一个t值, 在给定T=t条件下的条件联合分布只依赖t而不依赖. 也就是说，对于每一个t值, 在给定T=t条件下的条件分布对于任意值保持不变. 则我们称T是参数的充份统计量.

定理7.7.1 因式分解 设是连续分布或离散分布的随机样本，该分布的p.d.f.或p.f.为, 其中值是未知的并且属于给定的参数空间. 统计量是充分统计量当且仅当的联合p.d.f.或联合p.f. 对于所有的值和所有的值可以被分解称如下公式：

函数和是非负的，可能依赖但是不依赖，依赖于但是依赖通过统计量观察到的值.

推论7.7.1 统计量是充分的当且仅当无论使用哪种先验分布，后验分布只依赖通过T值的观察数据.

7.8 联合充分统计量 2019年9月5日10点12分

定义7.8.1 联合统计量 对于每一个和每一个的可能值, 假设在给定条件下的条件联合分布不依赖值. 则被称为的联合充分统计量.

定理7.8.1 联合充分统计量的因式分解 设是n个实变量函数. 统计量是的联合充分统计量当且仅当联合p.d.f.或联合p.f. 对于所有的值和所有的值可以被分解称如下公式：

定义7.8.2 有序统计量 假设是某个分布中的样本空间. 设是这些随机样本中最小值, 设是第二最小值，是第三最小值，一次类推. 在这种方式中, 表示样本中最大值, 表示下一个最大值. 随机变量被称为样本的有序统计量.

定理7.8.2 有序统计量在随机样本中是充分的 设是某个分布中的样本空间, 该分布p.d.f.或p.f.是. 则有序统计量对于是联合充分的.

定义7.8.3 极小（联合）充分统计量 统计量是极小充分统计量当且是充分的并且是其它充分统计量的函数. 统计向量是极小充分统计量当且是充分的并且是其它充分统计量的函数.

定理7.8.3 M.L.E和充分统计量 设对于是充分统计量. 则的M.L.E. 仅依赖通过统计量的观察数据.

定理7.8.4 贝叶斯估算器和充分统计量 设对于是充分统计量. 则的每一个贝叶斯估算器 仅依赖通过统计量的观察数据.