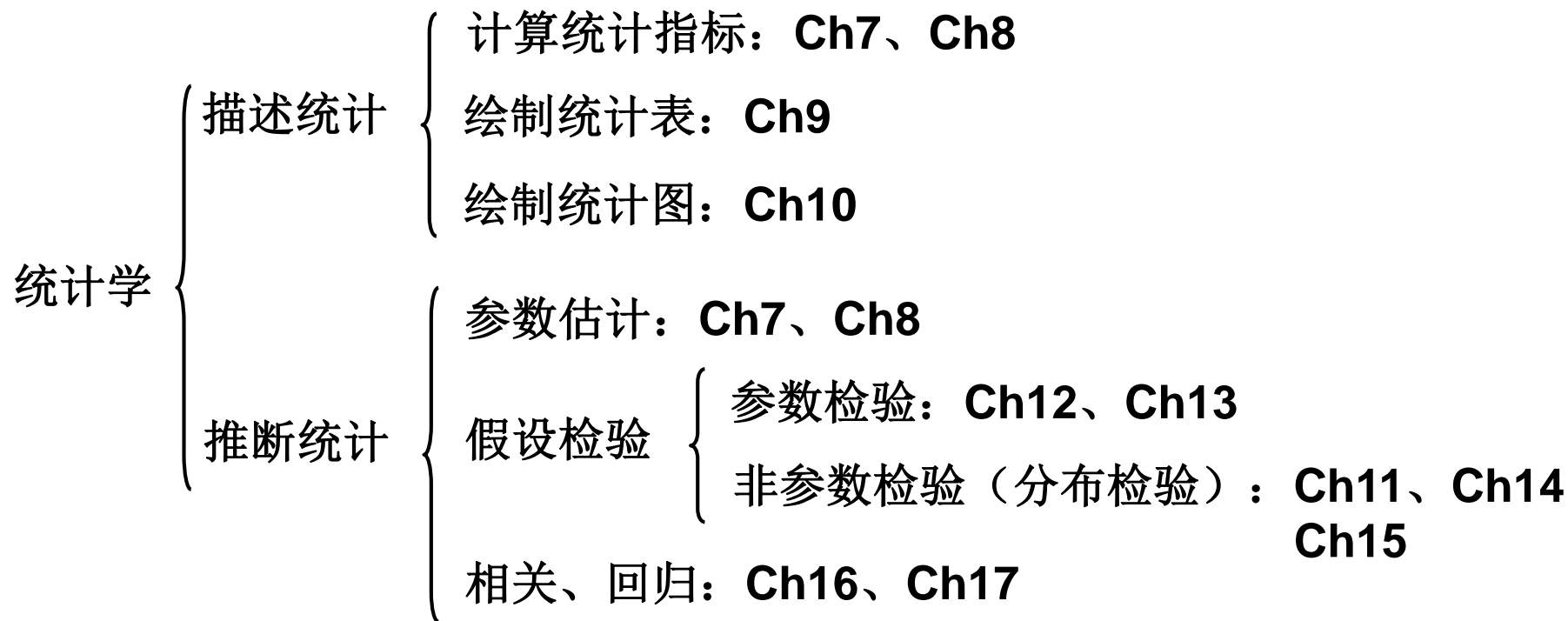




南京大學
NANJING UNIVERSITY

第11章 分布类型的检验



学习目标

- 掌握假设检验的基本思想
- 掌握如何对正态分布、二项分布、数据随机性进行检验

主要内容

- 11.1 假设检验的基本思想
- 11.2 正态分布检验
- 11.3 二项分布检验
- 11.4 游程检验（数据随机性检验）

11.1 假设检验的基本思想

生活中隐含的假设检验

- 掷骰子，猜到点数为胜

- 如果骰子没问题，则六个点的出现概率应当相等（均为 $1/6$ ，这就是一个事先假设），我们只是看每次具体的试验中谁的运气好。

- 今天一共下了600次注，竟然一共只猜中了一次

- 情况1：虽然平均应当赢约100次，但今天运气实在太差。

- 情况2：骰子有鬼，掷骰子的人可以人为控制结局，从而利用这种能力使自己得到了更多的收益。

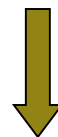
- 虽然第一种解释是可能的，但我们认为在骰子均匀的前提假设下出现如此结果实在是太不可能了（概率小到不应当被我们一次就碰上），因此我们认为骰子实际上不均匀。

生活中隐含的假设检验

假设：骰子均匀



大约赢**100**次
(而实际只赢了**1**次，产生矛盾)



假设不成立（拒绝假设）

注：假设骰子均匀，在这个前提条件下，投注**600**次中**1**次是一个小概率事件，而现在这个**小概率事件一次试验就发生了，这是一个矛盾**，产生矛盾的原因就是假设！

生活中隐含的假设检验

- 某学校某年级男生千米跑成绩均值为3分50秒，两个月前来了一名新的长跑教练，经过两个月的教学训练之后，从中随机抽测了10名男生的千米跑成绩，得到其样本均值为3分30秒，标准差为20秒，这时需要检验的问题是，新教练的训练方法是否使男生千米跑的成绩发生了改变？

生活中隐含的假设检验

假设：男生千米跑成绩没有发生改变，还是**3分50秒**



随机抽取**10个男生**，均值**3分30秒**，方差**20秒**，这是个**小概率事件**（小于等于**0.05**），而它现在在一次试验就发生了，产生了矛盾！



假设不成立（拒绝假设）

注：以上就是小概率反证法！

生活中隐含的假设检验

- 总结以上分析过程得到假设检验的流程：
 - 1.建立假设。 H_0 : 均值还是3分50秒; H_1 : 均值不是3分50秒。
 - 2.确定小概率事件的标准, $\alpha = 0.05$ 。
 - 3.计算在假设条件下样本事件发生的概率 P 。
 - 4.判断。如果 $P \leq \alpha$,说明一次试验就发生了小概率事件, 认为是矛盾的, 反推出 H_0 不成立; 如果 $P > \alpha$,说明一次试验出现了正常事件, 没有矛盾, 不能推翻 H_0 。

生活中隐含的假设检验

- 某工厂生产一种建筑材料——青砖，在原来的配方下，其抗压强度服从正态分布，现在改变了配料的方案。在改变配方之后，青砖的抗压强度是否还是正态分布呢？

H_0 : 改变配方后，抗压强度服从正态分布



从改变配方后的总体中抽取样本，看是否为小概率事件

拒绝假设或不拒绝假设

假设检验的基本思想

- **基础：小概率事件原理**，即一般认为小概率事件在一次随机抽样中不会发生。
 - 小概率事件：飞机失事
- **基本思想**：先建立一个关于样本所属总体的假设，考察在假设条件下随机样本的特征信息是否属小概率事件，若为小概率事件，则怀疑假设成立有悖于该样本所提供特征信息，因此拒绝假设。（**小概率反证法**）
- 事实上，小概率事件在随机抽样中还是可能发生的，只是发生的概率很小。若正好碰上了，则假设检验的结论就是错误的。当然，犯这种错误的概率很小。

假设检验的基本步骤

- 一、建立假设

- 根据统计推断的目的而提出的对总体特征的假设。统计学中的假设有两方面的内容：

- 一是**检验假设**(hypothesis to be tested)，也称**原假设、零假设或无效假设**(null hypothesis)，记为 H_0
 - 二是与 H_0 相对立的**备择假设**(alternative hypothesis)，记为 H_1 ，它的意义在于当 H_0 被拒绝时供采用。

假设检验的基本步骤

• 二、确立检验水准

- 确定检验水准（ α 水准，size of test）就是设立小概率事件的界值， α 取较小的值。
- 实际上 α 就是确定拒绝 H_0 时的最大允许误差的概率。是指检验假设 H_0 本来是成立的，而根据样本信息拒绝 H_0 的可能性大小的度量，换言之， α 是拒绝了实际上成立的 H_0 的概率上限。即 $P(\underbrace{\text{拒绝}H_0 | H_0\text{真}}_{\text{小概率事件}})$ 小于等于 α 。
- 常用的检验水准为 $\alpha = 0.05$ 。

假设检验的基本步骤

- 三、计算检验统计量和P值

- 进行试验，得到样本数据 X_1, X_2, \dots, X_n 。
- **统计量**：关于样本的函数 $F(X_1, \dots, X_n)$ 。
- 统计量只是工具，概率值才是目的，它可以客观**衡量样本对假设总体的偏离程度**。
 - **P值**：从 H_0 假设的总体中抽出现有样本（或更极端的情况）的概率。

假设检验的基本步骤

• 四、得出推断结论

– 按照事先确定的检验水准 α ，界定上面得到的P值，并按小概率事件原理认定对 H_0 的取舍，作出推断结论。

– 若 $P \leq \alpha$

- 在 H_0 假设下，一次试验就出现了小概率事件，则拒绝 H_0 ，接受 H_1 ，可以认为样本与总体的差别不仅仅是抽样误差造成的，可能存在本质上的差别，属“非偶然的（significant）”或显著的，因此，可以认为两者的差别有统计学意义。
- 进一步根据样本信息引申，得出实用性的信息。

假设检验的基本步骤

– 若 $P > \alpha$

- 基于 H_0 出现了很常见的事件。
- 则样本与总体间的差别尚不能排除纯粹由抽样误差造成，可能的确属“偶然的（non-significant）”或不显著的，故尚不能拒绝 H_0 （不拒绝并不代表接受，只是没有找到拒绝的理由），因此，认为两者的差别无统计学意义。

假设检验的两类错误

所作判断 真实情况	拒绝 H_0	不拒绝 H_0
H_0 为真	第一类错误 (弃真)	判断正确
H_0 为假	判断正确	第二类错误 (取伪)

犯第一类错误的概率通常记为 α (就是小概率事件临界值)

犯第二类错误的概率通常记为 β

假设检验的两类错误

- 任何检验方法都不能完全排除犯错误的可能性。
理想的检验方法应使犯两类错误的概率都很小，但在样本容量给定的情形下，不可能使两者都很小，降低一个，往往使另一个增大。
- 假设检验的指导思想是控制犯第一类错误的概率不超过 α ，然后，若有必要，通过增大样本容量的方法来降低 β 。

11.2 正态分布检验

预备检验

- 偏度系数、峰度系数
- 直方图（带分布曲线）
- P-P图、Q-Q图

K-S检验

- 是一种拟合优度检验
- 考察的是累积实际频率与累积理论频率的偏差。
 - 假设样本来自的总体服从正态分布。
 - 分别计算累积实际频率 (A_i) 和累积理论频率 (O_i) 。
 - 在假设条件下, 如果 A_i 和 O_i 相差较大, 这是一个小概率事件, 出现矛盾, 要拒绝假设; 反之, 如果 A_i 和 O_i 相差较小, 这是一个正常事件, 没有出现矛盾, 不能拒绝假设。
 - 由上述思路构造相应的统计量(服从K分布), 并计算样本事件发生的概率P, 根据P值与小概率事件标准 α 的比较来判断拒绝假设还是不拒绝假设。

K-S检验

- K-S检验可以检验的分布类型：
 - 正态分布
 - 均匀分布
 - 指数分布
 - 泊松分布

案例

- 数据文件：CCSS_Sample.sav
- 要求：检验2007年4月的变量index1是否服从正态分布
- 具体过程：
 - 选择个案
 - 计算偏度系数、峰度系数、绘制直方图、P-P图、Q-Q图
 - “分析” → “非参数检验” → “单样本”
 - “目标”选项卡：自动比较观察数据和假设数据
 - “字段”选项卡：选择“使用自定义字段分配”，将变量“index1”选入“检验字段”
 - “设置”选项卡：选择“自定义检验”中的K-S检验，并打开“选项”按钮选择“正态分布”

案例

 单样本非参数检验 ×

目标 字段 设置

使用一个或多个非参数检验识别单个字段中的差别。非参数检验不假定您的数据呈正态分布。

您的目标是什么？

每个目标对应“设置”选项卡上的一个不同缺省配置，如有需要，您可进一步自定义该配置。

☒ 自动比较观察数据和假设数据(U)

☐ 检验随机序列(S)

☐ 自定义分析(C)

描述

使用二项式检验、卡方检验或 Kolmogorov-Smirnov 自动比较观察数据和假设数据。 根据您的数据选择不同的检验。

 运行(R)  粘贴(P)  重置(R)  取消  帮助

案例

单样本非参数检验



目标 字段 设置

☐ 使用预定义角色(U)

☒ 使用自定义字段分配(C)

字段:

排序: 无(N)

A10. 那么与现在相比, 您认为五
A16. 对于大宗耐用消费品的购买
家庭收入2级
Qs9
Qa3
Qa4
Qa8
Qa9
Qa10
Qa16
现状指数
预期指数
time <= 200704 (FILTER)

全部



检验字段(I):

总指数



运行(R)

粘贴(P)

重置(R)

取消

帮助

案例

单样本非参数检验

目标 字段 设置

选择项目(S):

选择检验(S)

检验选项

用户缺失值

☐ 根据数据自动选择检验(U)

☒ 自定义检验(I)

☐ 比较观察二分类可能性和假

选项(B)...

☐ 比较观察可能性和假设可能

选项(P)...

☒ 检验观察分布和假设分布(K)

选项(K)...

☐ 比较中位数和假设中位数(I)

假设中位数(H):

☐ 检验随机序列(游程检验)(L)

选项...

Kolmogorov-Smirnov 检验选项

-假设分布-

☒ 正态(R)

分布参数

☒ 使用样本数据(S)

☐ 定制(U)

平均值(E): 0

标准差(V): 1

☐ 相等(F)

分布参数

☒ 使用样本数据(D)

☐ 定制(C)

最小值(I): 0

最大值(X): 1

☐ 指数分布(P)

平均值

☒ 样本平均值(L)

☐ 定制(M)

平均值(A): 0

☐ 泊松(O)

平均值

☒ 样本平均值(N)

☐ 定制(T)

平均值: 0

确定

取消

? 帮助

运行(R)

粘贴(P)

重置(R)

取消

? 帮助

案例

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of 总指数 is normal with mean 98.336 and standard deviation 18.92.	One-Sample Kolmogorov-Smirnov Test	.000 ¹	Reject the null hypothesis.

**K-S检验方法
的模型输出**

Asymptotic significances are displayed. The significance level is .05.

¹Lilliefors Corrected

- 分析结果

- P值 ≤ 0.05 , 拒绝原假设, 即认为数据不服从正态分布

- 补充说明

- 可以更改显著性水平 α (小概率事件的标准), 通过“设置”选项卡中的“检验选项”设置

11.3 二项分布检验

说明

- 是针对二分类变量的拟合优度检验
- 考察的是每个类别中观察值的频数与特定二项分布下的预期频数间是否存在统计学差异
 - 性别变量，检验男性（女性）的比例是否为 p

案例

- 数据文件: CCSS_Sample.sav
- 要求: 检验2007年4月份的男性比例是否为0.5
- 具体过程
 - 选择个案
 - “分析” → “非参数检验” → “单样本”
 - “目标”选项卡: 自动比较观察数据和假设数据
 - “字段”选项卡: 选择“使用自定义字段分配”, 将变量“性别”选入“检验字段”
 - “设置”选项卡: 选择“自定义检验”中的二项式检验, 并打开“选项”按钮将“假定比例”设为0.5

案例

 单样本非参数检验 ×

目标 字段 设置

使用一个或多个非参数检验识别单个字段中的差别。非参数检验不假定您的数据呈正态分布。

您的目标是什么？

每个目标对应“设置”选项卡上的一个不同缺省配置，如有需要，您可进一步自定义该配置。

☒ 自动比较观察数据和假设数据(U)

☐ 检验随机序列(S)

☐ 自定义分析(C)

描述

使用二项式检验、卡方检验或 Kolmogorov-Smirnov 自动比较观察数据和假设数据。 根据您的数据选择不同的检验。

 运行(R)  粘贴(P)  重置(R)  取消  帮助

案例



案例

单样本非参数检验



目标

字段

设置

选择项目(S):

选择检验(S)

检验选项

用户缺失值

☐ 根据数据自动选择检验(U)

☒ 自定义检验(I)

☒ 比较观察二分类可能性和假设二分类可能性(二项式检验)(O)

选项(B)...

☐ 比较观察可能性和假设可能性(卡方检验)(C)

选项(P)...

☐ 检验观察分布和假设分布(Kolmogorov-Smirnov 检验)(K)

选项(K)...

☐ 比较中位数和假设中位数(Wilcoxon 符号等级检验)(M)

假设中位数(H):

☐ 检验随机序列(游程检验)(Q)

选项...



运行(R)

粘贴(P)

重置(R)

取消

帮助

案例

二项式选项



假设比例(H): 0.5

置信区间

☐ Clopper-Pearson (精确)

☐ Jeffreys

☐ 似然比(L)

定义分类字段的成功

☒ 使用数据中找到的第一个类别(U)



☐ 指定成功值(S)

成功值(V):

值



定义连续字段的成功值



成功值等于或小于

☒ 样本中点(E)

☐ 定制分割点(I)

分割点(C):

确定

取消

? 帮助

案例

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The categories defined by S2. 性别 = 男 and 女 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.094	Retain the null hypothesis.

二项分布检验
的模型输出

Asymptotic significances are displayed. The significance level is .05.

- 分析结果

- $P \text{ 值 } \geq 0.05$, 不拒绝原假设, 即认为男性比例为0.5

11.4 游程检验

说明

- 针对二分类变量的随机检验
- 用于检验数据是否是随机抽取的
- 说明：连续变量也能进行游程检验，会事先将其分为两类

案例

- 数据文件：CCSS_Sample.sav
- 要求：检验2007年4月份的性别数据和年龄数据是否随机
- 具体过程
 - 选择个案，按ID升序排列
 - “分析” → “非参数检验” → “单样本”
 - “目标”选项卡：检验随机序列
 - “字段”选项卡：选择“使用自定义字段分配”，将变量“性别”和“年龄”选入“检验字段”
 - “设置”选项卡：选择“自定义检验”中的游程检验

案例

 单样本非参数检验 ×

目标

字段

设置

使用一个或多个非参数检验识别单个字段中的差别。非参数检验不假定您的数据呈正态分布。

您的目标是什么？

每个目标对应“设置”选项卡上的一个不同缺省配置，如有需要，您可进一步自定义该配置。

☐ 自动比较观察数据和假设数据(U)

☒ 检验随机序列(S)

☐ 自定义分析(C)

描述

使用游程检验检验随机序列。

▶ 运行(R)

粘贴(P)

重置(R)

取消

🔍 帮助

案例

单样本非参数检验



目标

字段

设置

☐ 使用预定义角色(U)

☒ 使用自定义字段分配(C)

字段:

排序: 无(N)

- 月份
- ID
- S0. 城市
- S4. 学历
- S5. 职业
- S7. 婚姻状况
- S9. 家庭月收入
- C0. 请问您的家庭目前有下列还
- C0. 请问您的家庭目前有下列还
- C0. 请问您的家庭目前有下列还
- O1. 是否拥有家用轿车
- A3. 首先, 请问与一年前相比, 您
- A3a. 您为什么这么说呢?

全部

检验字段(T):

- S2. 性别
- S3. 年龄

运行(R)

粘贴(P)

重置(R)

取消

帮助

案例

单样本非参数检验

目标 字段 设置

选择项目(S):

选择检验(S)

检验选项

用户缺失值

☐ 根据数据自动选择检验(U)

☒ 自定义检验(I)

☐ 比较观察二分类可能性和假设二分类可能性

选项(B)...

☐ 比较观察可能性和假设可能性 (卡方检验)

选项(P)...

☐ 检验观察分布和假设分布 (Kolmogorov-S

选项(K)...

☐ 比较中位数和假设中位数 (Wilcoxon 符号

假设中位数(H):

☒ 检验随机序列 (游程检验) (Q)

选项...

游程检验选项

定义分类字段的组

☒ 样本中仅有 2 个类别(M)

☐ 将数据重新编码为 2 个类别(R)

定义第一个类别(D):

值

定义连续字段的分割点

☒ 样本中位数(S)

☐ 样本平均值(E)

☐ 定制(U)

分割点(C):

确定

取消

? 帮助

运行(R)

粘贴(P)

重置(R)

取消

? 帮助

案例

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The sequence of values defined by S2. 性别 = (男) and (女) is random.	One-Sample Runs Test	.683	Retain the null hypothesis.
2	The sequence of values defined by S3. 年龄 ≤ 35.000 and > 35.000 is random.	One-Sample Runs Test	.808	Retain the null hypothesis.

游程检验
的模型输出

Asymptotic significances are displayed. The significance level is .05.

- 分析结果

- 两个P值均 ≥ 0.05 , 不拒绝原假设, 即认为性别数据和年龄数据均是随机的

- 补充说明

- 针对连续变量, 分为两类的分割点可以选择: 中位数、平均值、自定义值

THE END