

# Lecture 6 – 假设检验 Hypothesis testing

- 内容大纲
  - 知识回顾 L01-L05: statistics and uncertainty
  - 假设检验 Hypothesis testing
  - I类和II类错误 Type I and Type II errors
  - 单侧和双侧检验 One-sided and two-sided tests
  - 课堂练习 Practice Problems
  - 小结 Summary

生物统计学

李 勤

生态与环境科学学院

# 第二次课后作业

- 作业内容:
  - Lecture 4-6
  - 统计量及其不确定性、概率计算、假设检验
- 作业发布时间：2024年11月1日
- 作业提交截止时间：2024年11月8日

# 1. 回顾 L01 – L05

- 统计学的目的是基于从总体中的样本来获得信息
  - 通过随机抽样获取样本（这也是我们做统计推断的常用假设）
  - 样本即子集：总体包含的每个个体有相同概率被抽中，且相互独立
  - 抽样误差：因随机抽样造成的估计值与真实值（总体参数）之间的偶然差异
- 基于样本应用不同估计值对总体进行推断（参数估计）
  - 均值、标准差、中值、四分位距
- 同时提供推断估计值的精确性
  - 标准误、置信区间（2SE rule）

# 1. 回顾 L01 – L05

- 概率 probability

- 对总体进行随机取样即是一种随机试验，其结果受概率规则的制约；
- 事件的概率是指在相同条件下反复进行随机试验时，事件发生的次数比例；
- 概率分布描述了随机试验所有可能结果的概率；

- 概率运算

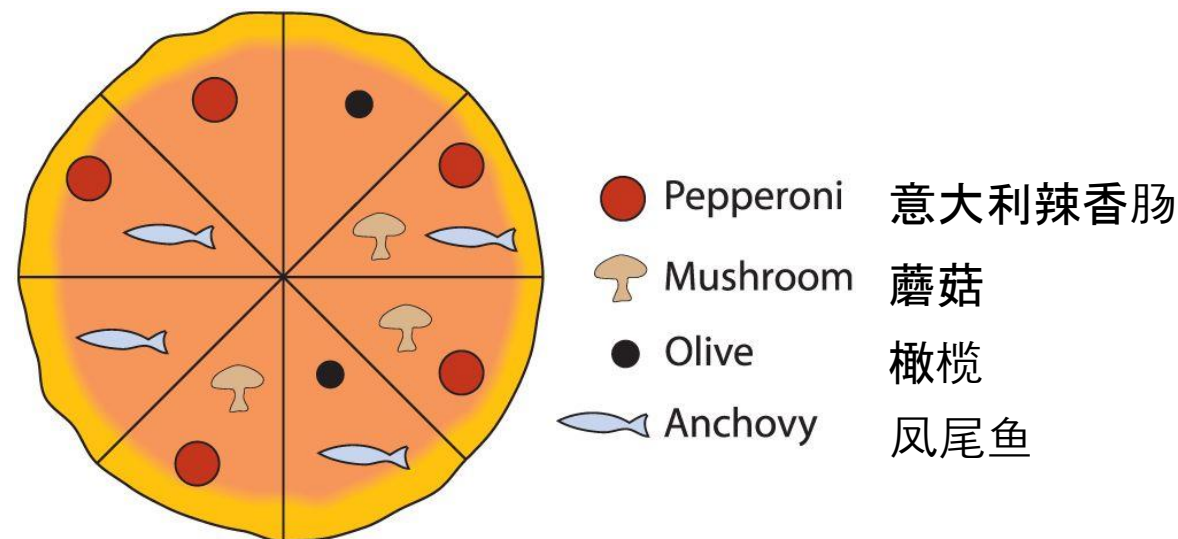
- 互斥事件参考加法公式： $\Pr[A \text{ or } B] = \Pr[A] + \Pr[B]$
- 扩展的一般加法公式为  $\Pr[A \text{ or } B] = \Pr[A] + \Pr[B] - \Pr[A \& B]$
- 独立事件参考乘法公式： $\Pr[A \& B] = \Pr[A] \times \Pr[B]$



# 1. 回顾 — 课堂讨论

- 外卖的披萨，分为大小相等的8片；每片披萨上可能有意大利香肠、蘑菇、橄榄、或凤尾鱼。试想一下，随机拿起一片披萨（即拿起八片披萨中任何一片的概率为  $1/8$ ）；

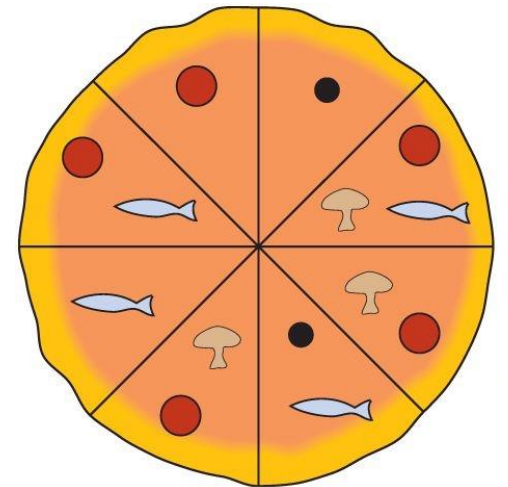
- 请根据今天的概率知识回答：



Whitlock & Schluter, *The Analysis of Biological Data*, 3e  
© 2020 W. H. Freeman and Company

# 1. 回顾 — 课堂讨论

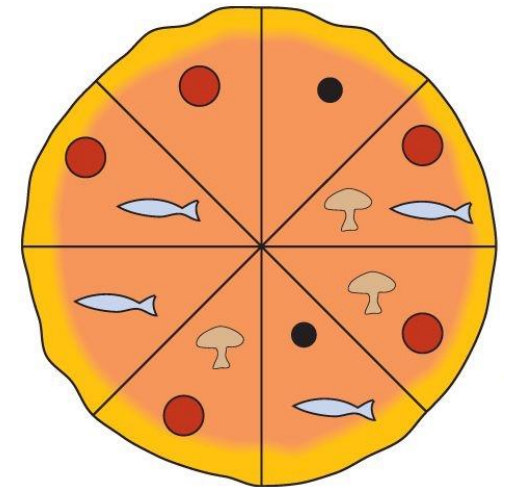
- 1. 你的披萨上有意大利辣香肠的概率是多少?
  - $\Pr[\text{pepperoni}] = 5/8$
- 2. 你的披萨上既有意大利辣香肠又有凤尾鱼的概率是多少?
  - $\Pr[\text{pepperoni} \ \& \ \text{anchovy}] = 2/8 = 1/4$
- 3. 你的披萨上有意大利辣香肠或凤尾鱼的概率是多少?
  - $\Pr[\text{pepperoni} \mid \text{anchovy}] = 7/8$
- 4. 这块比萨上的意大利辣香肠和凤尾鱼是互斥的吗?
  - 不是；有的披萨上既有意大利辣香肠又有凤尾鱼
- 5. 这块比萨上的橄榄和蘑菇是互斥的吗?
  - 是，没有一片披萨上既有橄榄又有蘑菇



# 1. 回顾 — 课堂讨论

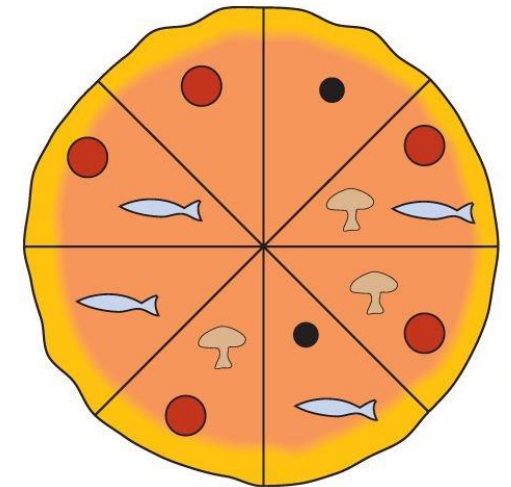
- 6. 在选择一片比萨时，蘑菇和风尾鱼是独立的吗？
  - $\Pr[\text{mushrooms}] = 3/8$ ;  $\Pr[\text{anchovies}] = 4/8 = 1/2$ ;
  - 如果独立，那么  $\Pr[\text{mushrooms} \& \text{anchovies}] = \Pr[\text{mushrooms}] \times \Pr[\text{anchovies}] = 3/16$
  - 但实际上，  $\Pr[\text{mushrooms} \& \text{anchovies}] = 1/8$
  - 所以，蘑菇和风尾鱼不是独立的
- 7. 如果我从这个比萨中挑选一片，并告诉你上面有橄榄，那么它也有风尾鱼的概率有多大？
  - 条件概率：  $\Pr[\text{olives} | \text{anchovies}] = 1/2$
- 8. 如果我从这块披萨中挑选一块，并告诉您上面有风尾鱼，那么它也有橄榄的概率是多少？
  - 条件概率：  $\Pr[\text{olives} | \text{anchovies}] = 1/4$

	Pepperoni	意大利辣香肠
	Mushroom	蘑菇
	Olive	橄榄
	Anchovy	凤尾鱼



# 1. 回顾 — 课堂讨论

- 9. 你的七个朋友每人随机选择一片披萨，吃完后不告诉你他们吃了什么配料。那么，剩下的最后一块有橄榄的概率是多少？
  - $\Pr[\text{last slice has olives}] = \Pr[\text{olive}] = 2/8 = 1/4$
- 10. 你随机选择两片披萨，它们都有橄榄的概率是多少？
  - (提示：选掉第一片后，从剩下的披萨片中选择一片的概率会发生变化)
  - $\Pr[\text{two slices with olives}] = \Pr[\text{first slice has olives}] \times \Pr[\text{second slice has olives} | \text{first slice has olives}] = 2/8 \times 1/7 = 1/28$



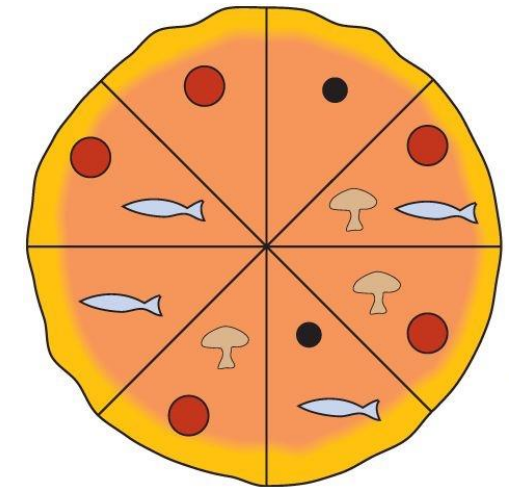


# 1. 回顾 — 课堂讨论

- 11. 随机选择的一片披萨上没有意大利辣香肠的概率是多少？

- $\Pr[\text{slice without pepperoni}] = 1 - \Pr[\text{slice with pepperoni}]$   
 $= 1 - 5/8 = 3/8$

- 12. 画一个同样分为8片的，且蘑菇、橄榄、凤尾鱼和意大利辣香肠相互排斥的披萨。



## 2. 假设检验 hypothesis testing

- 学习目标

- 理解假设检验的统计意义及其步骤;
- 掌握假设检验的步骤并运用其来尝试解释生物学现象;
- 熟知假设检验中的错误及其应对方式;

- 这节课会非常多的**术语**



## 2. 假设检验 hypothesis testing

- 统计学：基于样本来对总体进行推断； 并做假设检验；

- mean
- median
- SD, IQR
- proportion
- SE, CI

参数估计  
parameter  
estimation

为总体特征给出数值及范围  
(values and bounds)

“新药的治疗效果有多大？”  
“How large is the effect?”

假设检验  
hypothesis  
testing

$H_0$

$H_A$

询问参数是否不同于特定的期望  
(a specific “null” expectation  $H_0$ )

“新药有治疗效果吗？”  
“Is there any effect at all?”

## 2.1 提出并使用统计假设

- 假设检验
  - 设定零假设
  - 将实际数据与我们在特定零假设为真时的期望结果进行比较
  - 如果数据与零假设成立时的预期数据相比过于异常，则拒绝零假设
- 对单一总体中的比例进行检验
  - 以比例为例来说明假设检验的基本原理（最简单的示范）
    - 目标：最少的计算
  - 本课之后的内容将介绍假设检验的其它许多具体方法

## 2.1 提出并使用统计假设

- 统计假设 vs 科学假设
  - 统计假设：包括零假设和备择假设都是关于总体特征的简单陈述
  - 科学假设：关于自然现象的存在和可能原因的陈述
  - **不要混淆统计假设与科学假设**
- 科学家设计实验研究和观察研究来检验科学假设的预测
  - 将统计假设应用于所得数据有助于确定这些科学假设的哪些预测与数据一致，哪些不一致。

## 2.2 零假设 A null hypothesis - $H_0$

- 零假设是关于总体参数值的一种具体说法
  - 它是为了论证的目的而提出的，通常是一个可被质疑的观点
  - 通常，零假设是指我们感兴趣的总体参数为零，
    - 即无影响、无偏好、无相关、或无差异
    - i.e., no effect, no preference, no correlation, or no difference
- 一般来说，零效假设是人们倾向于拒绝的一个陈述
  - 特定的或基于理论和已有知识预期

## 2.3 备择假设 An alternative hypothesis -H<sub>A</sub>

- 备择假设包括零假设中所述值之外的所有其它可能的总体参数值
  - 通常包括在生物学上比零假设中的值更有趣的可能性 (biologically more interesting)。
- 备择假设通常包括科学假设所预测的参数值
  - 因此研究人员通常希望备择假设为真（但并非总是如此）。
- 与零假设相比，备择假设是非特异性的
  - 除了零假设中特定值以外，总体特征参数的所有可能值或任何对比都包括在内。
  - 一样 vs. 不一样（或大或小）

## 2.4 零假设和备择假设

- 例如：

- 一项调查流网捕鱼 (drift-net fishing) 对海豚密度影响的研究

- $H_0$ : 在有流网捕鱼的区域和没有流网捕鱼的区域, 海豚密度相同。

(The density of dolphins is the same in areas with and without drift-net fishing.)

- $H_A$ : 在有流网捕鱼的区域和没有流网捕鱼的区域, 海豚密度不相同。

(The density of dolphins *differs* between areas with and without drift-net fishing.)

- $H_0$ : 一对父母眼睛为棕色, 他们各有一个蓝色眼睛的父母, 那么, 他们子女中棕色和蓝色眼睛的比例为 3:1。

- $H_A$ : 一对父母眼睛为棕色, 他们各有一个蓝色眼睛的父母, 那么, 他们子女中棕色和蓝色眼睛的比例是 3:1以外的比例。

- $H_0$ : 健康人的平均体温为37 摄氏度。

- $H_A$ : 健康人的平均体温不是37 摄氏度。



## 2.5 假设检验: 示例——handedness in toads

- 假设检验的步骤:
  - 1. 提出假设 (state the  $H_0$  &  $H_A$ )
  - 2. 利用数据计算检验统计量 (compute the test statistic with data)
  - 3. 确定 P 值 (determine the P-value)
  - 4. 得出适当的结论 (draw the appropriate conclusions)
  - 5. 报告结果 (Reporting the results)

## 2.5 假设检验: 示例——handedness in toads

- 假设检验的5个步骤
- 示例: European toads (*Bufo bufo*) (Bisazza et al. 1996)
  - 人类主要是右撇子。
  - 其它动物是否也有惯用手的现象?
- 注意:
  - 这个示例针对的是**比例**——以说明基本过程
  - 假设检验可以应用于多种统计量, 如均值、方差、均值差异、相关性等
  - 在未来的学习中会讨论有关其它统计量计算的细节



Hintau Aliaksei/Shutterstock.com

## 2.5 假设检验: 示例——handedness in toads

- 步骤一: 提出假设

- 怎么提出 (或确定) 零假设 vs 备择假设?

- 科学问题 (科学假设) :

- 在蟾蜍总体中右撇子(right-handed)和左撇子(left-handed)出现的频数相同, 还是像人类那样一种类型比另一种类型更常见?

- 统计假设:

- $H_0$ : 频数相同

$$H_0: p_{left} = p_{right} = 0.5$$

- $H_A$ : 频数不同

$$H_A: p_{left} \neq p_{right}$$

- 双边检验 two-sided test

- $p_{left} > p_{right}$  or  $p_{left} < p_{right}$

- 在获取数据前无法排除任何一种可能性 (方向)

## 2.5 假设检验: 示例——handedness in toads

- 步骤一: 提出假设

- $H_0$ : 频数相同
- $H_A$ : 频数不同

$$H_0: p_{\text{left}} = p_{\text{right}} = 0.5$$

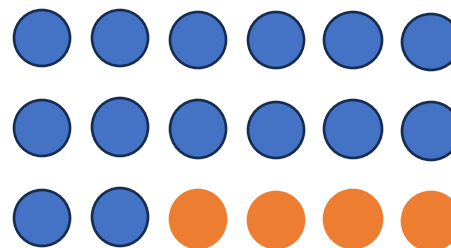
$$H_A: p_{\text{left}} \neq p_{\text{right}}$$

- 实验体系:

- 样本: 野外捕获的18只蟾蜍 (假设这是一个随机样本)
- 测试方法:
  - 研究人员给每只蟾蜍头上都缠着一个气球
  - 记录下每只蟾蜍用哪只前肢去除掉气球

- 实验结果:

- 14 right-handed
- 4 left-handed



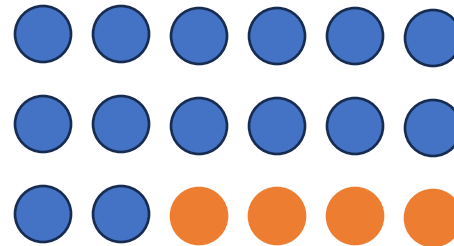
Hintau Aliaksei/Shutterstock.com



## 2.5 假设检验: 示例——handedness in toads

- 步骤二: 根据数据计算检验统计量
  - 检验统计量是根据数据计算出来的数值
    - 用以评估实际数据的观察结果与零假设 $H_0$ 条件下的预期结果的吻合程度
- 这里的统计量为: 比例
  - 表征18只蟾蜍中使用右前肢的偏好 (右撇子): 14/18

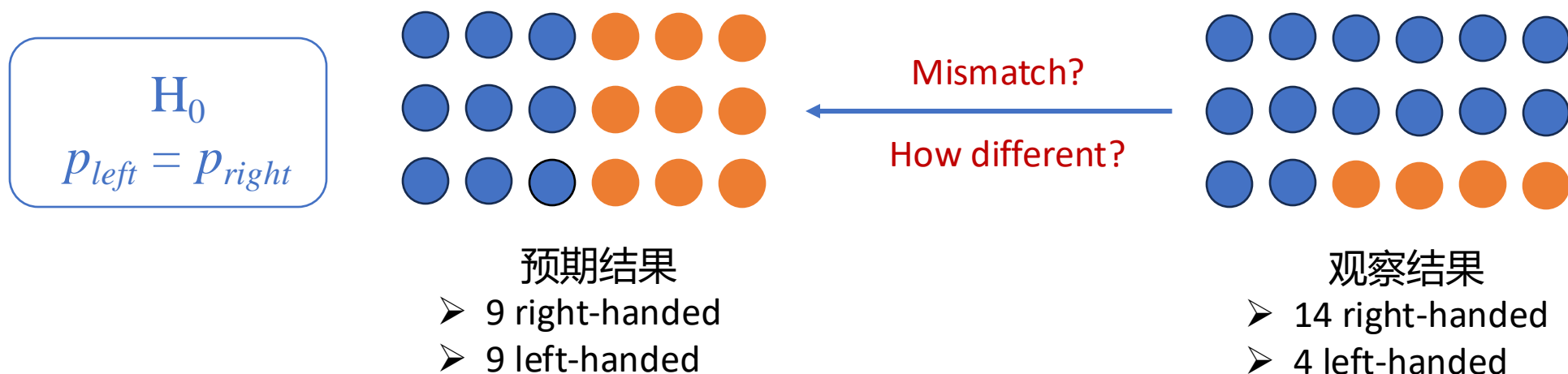
- 14 right-handed
- 4 left-handed



- 这个结果是否说明蟾蜍有手性偏好呢?

## 2.5 假设检验: 示例——handedness in toads

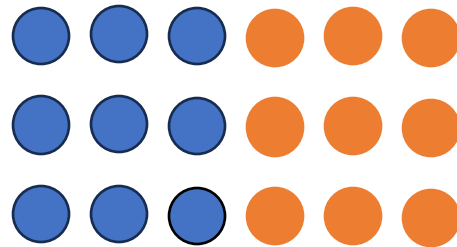
- 步骤三: 确定 P 值 (量化不确定性)
  - 检验统计量用以评估观察结果与预期结果的吻合程度
    - 由于抽样过程中偶然性的影响 (特别是在数据不多的情况下), 观察结果与 $H_0$ 下的预期结果之间通常会存在差异;
    - 所以, 要判断数据是否符合 $H_0$ , 我们必须计算出当假设零假设成立 (为真)时, 出现与观察结果一样或比观察结果更极端的情况的概率。



## 2.5 假设检验: 示例——handedness in toads

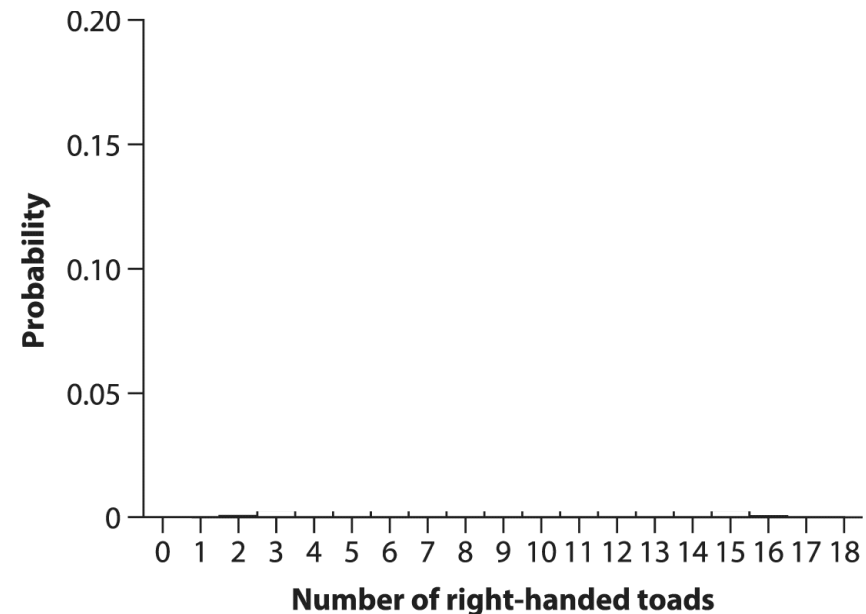
- 步骤三: 确定 P 值 (量化不确定性)
  - 要获得这个概率, 我们需要确定当 $H_0$ 为真时的检验统计量的抽样分布:
    - $H_0$ 下检验统计量的可能值及其相关概率 → 零分布 (the null distribution)
    - 可利用计算机的高性能来进行计算

$$H_0$$
$$p_{\text{left}} = p_{\text{right}}$$



预期结果

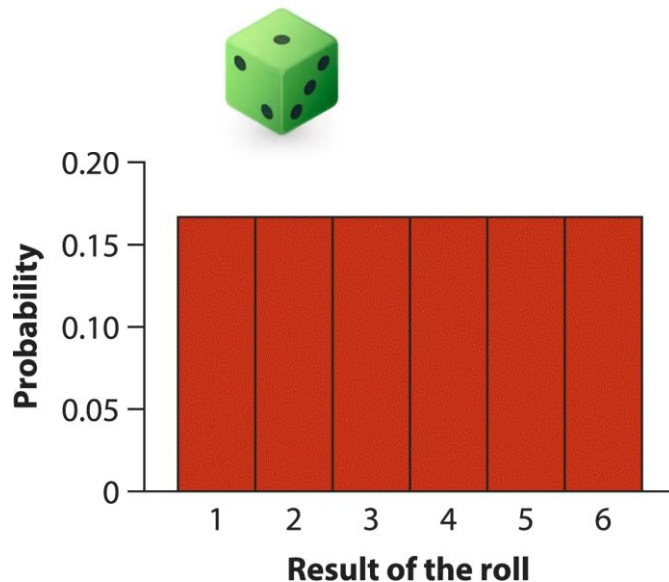
- 9 right-handed
- 9 left-handed



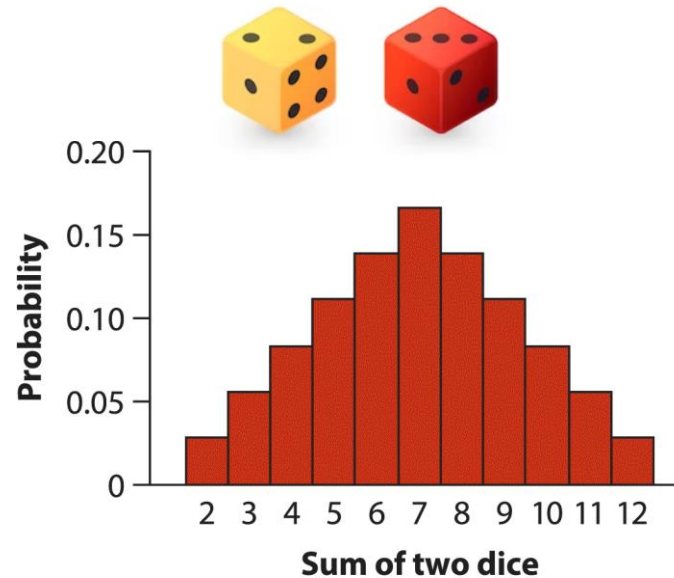
Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

## 2.5 假设检验: 示例——handedness in toads

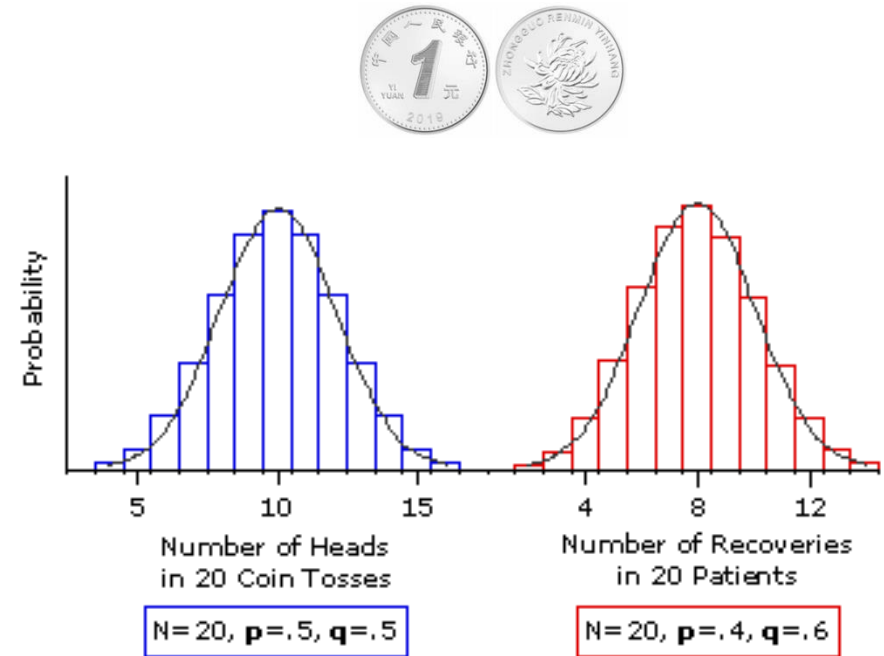
- 步骤三: 确定 P 值 (量化不确定性)
  - 检验统计量的抽样分布——零分布
  - 同等条件下大量重复试验的概率分布 (包含各种可能结果)



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company



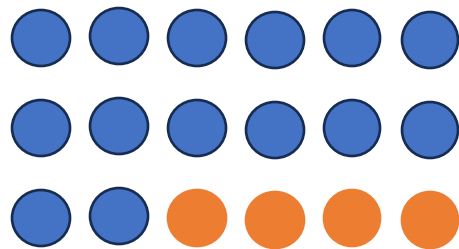
Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company





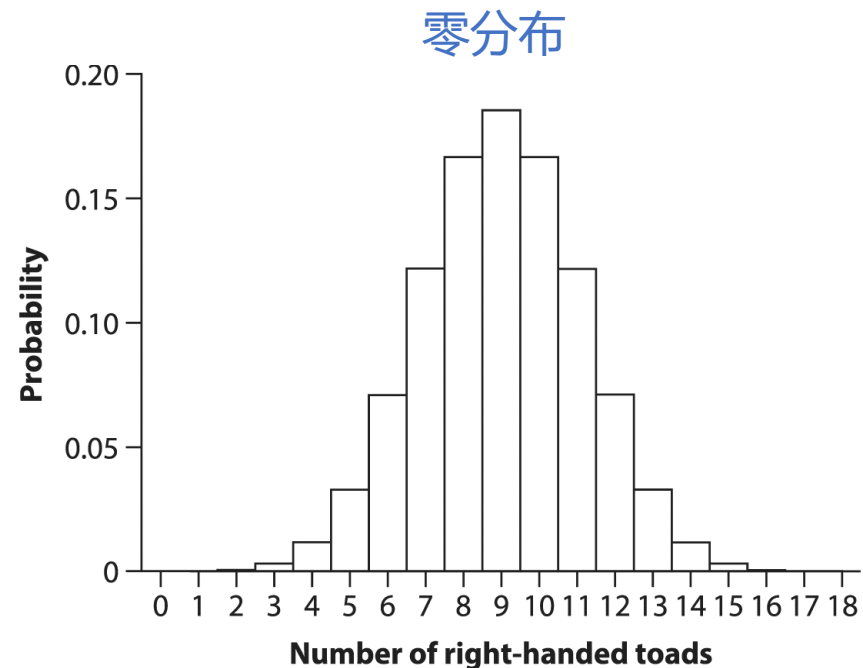
## 2.5 假设检验: 示例——handedness in toads

- 步骤三: 确定 P 值 (量化不确定性)
  - 检验统计量的抽样分布——零分布
    - 18只蟾蜍中使用右前肢的偏好 (右撇子): 比例
    - 类似于抛18枚硬币, 然后记录其中数字朝上 (两种结果之一) 的个数及其比例
    - 18只蟾蜍中右撇子的个数  $\subset [0, 18]$



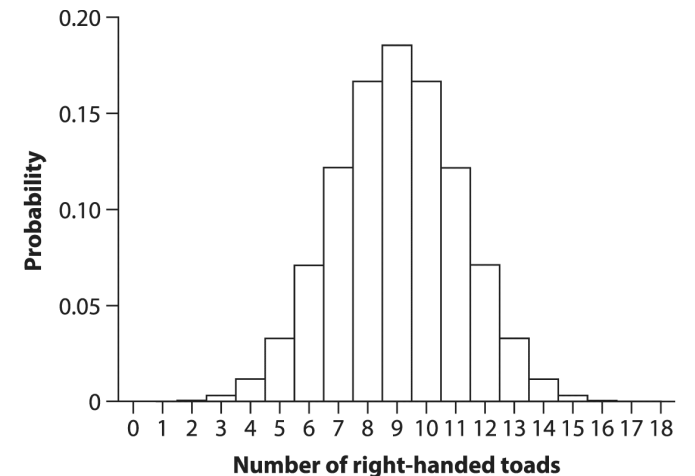
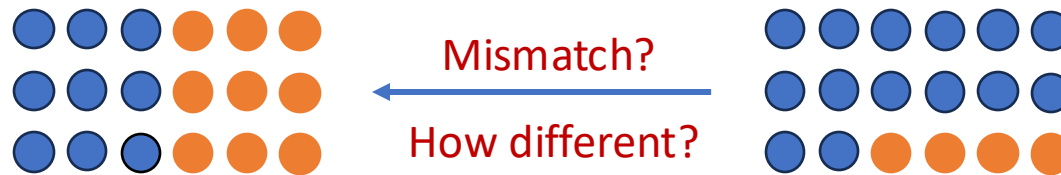
观察结果

- 14 right-handed
- 4 left-handed



## 2.5 假设检验: 示例——handedness in toads

- 步骤三: 确定 P 值 (量化不确定性)
  - $14/18 \neq 9/18$ ?
  - 但这种不匹配是否大到足以排除偶然性?
    - 要判断数据是否符合 $H_0$ , 我们必须计算出当假设零假设为真时, 出现与观察结果一样或比观察结果更极端的情况的概率?
- $\Pr[\text{大于或等于}14\text{只蟾蜍是右撇子}] = ?$

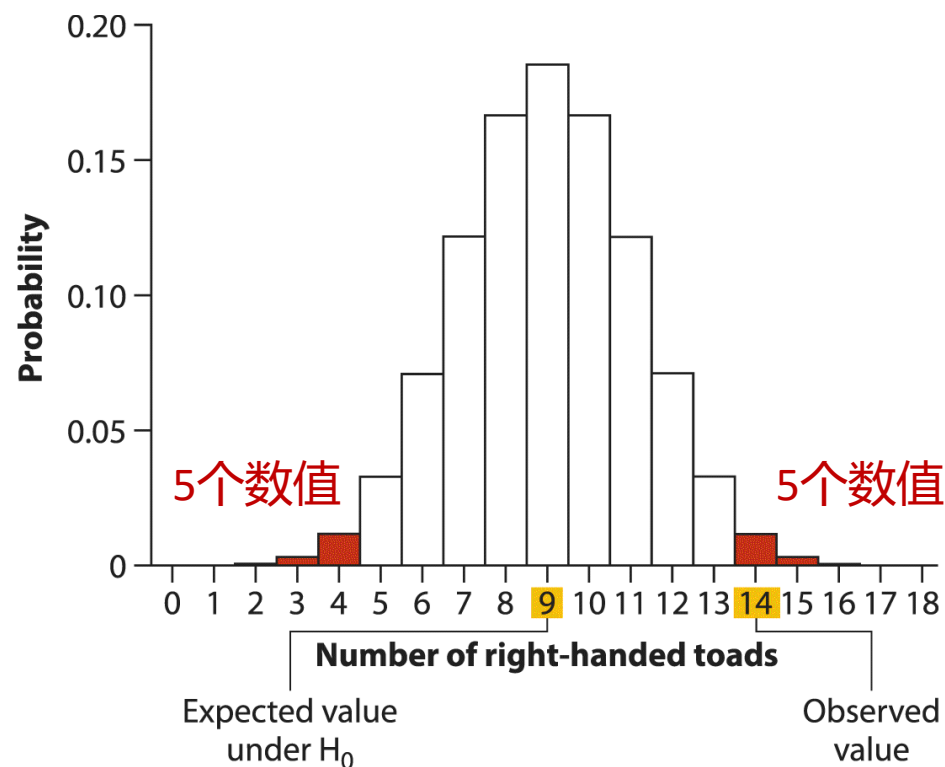


Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

## 2.5 假设检验: 示例——handedness in toads

- 步骤三: 确定 P 值 (量化不确定性)
  - 计算 P 值
    - 需要考虑到与观测数据一样不寻常或甚至更极端的结果:
    - 双边检测则需要考虑:
      - 右撇子更多
        - $\Pr[\text{大于或等于14只蟾蜍是右撇子}]$
      - 左撇子更多
        - $\Pr[\text{大于或等于14只蟾蜍是左撇子}]$   
 $= \Pr[\text{小于或等于4只蟾蜍是右撇子}]$
        - 4只右撇子 = 14只左撇子
    - 即零分布的左右尾部的各5 个数字

$H_0$ : 左撇子和右撇子频数相同



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

## 2.5 假设检验: 示例——handedness in toads

- 步骤三: 确定 P 值 (量化不确定性)

- 假设零假设 $H_0$ 为真的前提下, 计算P 值

(两侧是对称的)

- $\Pr[\geq 14 \text{ right-handed toads}] = \Pr[14] + \Pr[15] + \Pr[16] + \Pr[17] + \Pr[18]$

$$= 0.0155$$

- $\Pr[\geq 14 \text{ left-handed toads}] = \Pr[4] + \Pr[3] + \Pr[2] + \Pr[1] + \Pr[0]$

$$= 0.0155$$

- $\text{P-value} = \Pr[\geq 14 \text{ right-handed toads}] + \Pr[\geq 14 \text{ left-handed toads}]$

$$= 2 * 0.0155 = \mathbf{0.031}$$

Number of right-handed toads	Probability
0	0.000004
1	0.00007
2	0.0006
3	0.0031
4	0.0117
5	0.0327
6	0.0708
7	0.1214
8	0.1669
9	0.1855
10	0.1669
11	0.1214
12	0.0708
13	0.0327
14	0.0117
15	0.0031
16	0.0006
17	0.00007
18	0.000004
Total	1

## 2.5 假设检验: 示例——handedness in toads

- 步骤三: 确定 P 值 (量化不确定性)
  - P 值: 在零假设 $H_0$ 为真的情况下获得观察数据 (或更极端结果) 的概率
    - 如果 P 值很小, 那么意味着数据与 $H_0$ 的预期结果不一致, 我们需要拒绝 $H_0$ ;
    - 否则, 我们不拒绝 $H_0$ ;
    - 一般来说, P 值越小, 反对 $H_0$ 的证据就越强;
  - P 值**不是**零假设为真的概率
    - “假设” (hypothesis) 不是随机试验的结果, 因此没有概率;
    - P 值指的是在零假设条件下进行数据抽样时发生特定事件的概率;



## 2.5 假设检验: 示例——handedness in toads

- 步骤四: 给出恰当的结论
  - 如果  $P$  值很 "小", 则拒绝零假设; 否则, 不拒绝  $H_0$ .
    - $P$  值 = 0.031
    - 因此, 我们需要某种标准 (临界值/阈值) 来做判断;
  - $P$  值多小才算足够小?
    - 这个 临界值称为显著性水平, 用  $\alpha$  表示;
    - 按照大多数生物研究领域的惯例,  $P$  值的临界值  $\alpha = 0.05$ ;
    - 如果  $P \leq 0.05$ , 我们就拒绝零假设; 如果  $P > 0.05$ , 我们就不拒绝它;
- 那么, 蟾蜍有手性偏好吗?



## 2.5 假设检验: 示例——handedness in toads

- 步骤四: 给出恰当的结论

- 如果 P 值很 "小", 则拒绝零假设; 否则, 不拒绝  $H_0$ 。

- 那么, 蟾蜍有手性偏好吗?

- 统计假设的结论:

- 在  $\alpha = 0.05$  的前提下, P 值 = **0.031 是显著的** (significant) ;
    - 我们拒绝在蟾蜍总体中左撇子和右撇子出现频率相同的零假设;

- 科学假设的结论:

- 我们基于已有数据得出的结论是: 大多数蟾蜍是右撇子;



Hintau Aliaksei/Shutterstock.com

$$H_0: p_{\text{left}} = p_{\text{right}}$$



$$H_A: p_{\text{left}} < p_{\text{right}}$$

## 2.5 假设检验: 示例——handedness in toads

- 步骤五: 报告结果

- 在研究论文或实验报告中撰写统计检验结果时, 一般要包含以下信息:

- 检验统计量的值
    - 样本量
    - P 值

- 此外, 最好能提供相关统计参数的置信区间或标准误差。

- P 值表示反对零假设的证据的权重 (P 越小表示证据越有力)
    - 置信区间则衡量效应的大小

- 例如:

- 18 只蟾蜍中有 14 只是右撇子 (在这种情况下, 既给出了检验统计量, 也给出了样本量)
    - $P = 0.031$ ;  $0.54 < \text{Proportion} < 0.91$  (95% CI); 表明大多数蟾蜍是右撇子;



## 2.5 假设检验: 示例——handedness in toads

- 回顾——假设检验的步骤:
  - 1. 提出假设 (state the  $H_0$  &  $H_A$ )
  - 2. 利用数据计算检验统计量 (compute the test statistic with data)
  - 3. 确定 P 值 (determine the P-value)
  - 4. 得出适当的结论 (draw the appropriate conclusions)
  - 5. 报告结果 ( Reporting the results)

## 2.6 零假设和备择假设的区别

- 零假设和备择假设的地位并不相同，零假设是用数据进行检验的唯一陈述
  - 我们假设  $H_0$  为真，而假设检验是要看数据告诉我们什么信息；
  - 如果数据与零假设一致，意味着我们没有拒绝零假设(failed to reject it);
    - 我们从不“接受”零假设 (we never “accept” the null hypothesis) ;
  - 如果数据与零假设不一致，那么拒绝零假设，并说数据支持备择假设；
- 拒绝  $H_0$  意味着我们排除了零假设的值
  - 它还告诉我们，与零假设值相比，真实值可能的方向；
  - 但是，拒绝假设本身并不能揭示群体参数的大小或处理间差异的大小；
  - 我们使用估计值本身来提供大小；

### 3. 假设检验中的错误

- I 类错误和 II 类错误 Type I and Type II errors

- I 类错误：拒绝为真的零假设

- 显著性水平  $\alpha$  (the significance level) 设定了犯一类错误的概率；

- 生物学研究中通常将  $\alpha = 0.05$  视为可接受的错误率；

- 降低  $\alpha$  (如  $\alpha=0.01$ ) 会使得  $P > \alpha$ , 这样会增加拒绝  $H_0$  的难度；

- 好处：当  $H_0$  为真时，没有拒绝  $H_0$ ；

- 坏处：当  $H_0$  为假时，犯 II 类错误；

		Actual Situation	
		$H_0$ True	$H_0$ False
Decision	Fail to Reject $H_0$	✓	Type II Error ( $\beta$ )
	Reject $H_0$	Type I Error ( $\alpha$ )	✓

### 3. 假设检验中的错误

- I 类错误和 II 类错误 Type I and Type II errors
  - I 类错误：拒绝为真的零假设
  - II 类错误：没有拒绝为假的零假设
    - 如果一项研究犯 II 类错误的概率较低，我们则说它功效 (power) 较高；
    - 功效 (power)：拒绝为假的零假设的概率；
    - 在其它条件相同的情况下，一项研究的功效越大，则是越好的研究；

		Actual Situation	
		$H_0$ True	$H_0$ False
Decision	Fail to Reject $H_0$	✓	Type II Error ( $\beta$ )
	Reject $H_0$	Type I Error ( $\alpha$ )	✓

Power = 1 -  $\beta$

### 3. 假设检验中的错误

- I 类错误和 II 类错误 Type I and Type II errors
  - I 类错误：拒绝为真的零假设 —— **假阳性**
  - II 类错误：没有拒绝为假的零假设 —— **假阴性**

		Actual Situation	
		$H_0$ True	$H_0$ False
Decision	Fail to Reject $H_0$	✓	Type II Error ( $\beta$ )
	Reject $H_0$	Type I Error ( $\alpha$ )	✓

$H_0$ : 怀孕 (成立)



$H_0$ : 没怀孕 (不成立)



$H_0$ : 没怀孕 (成立)



$H_0$ : 怀孕 (不成立)



## 4. 假设检验中的单侧或双侧检验

- 单侧检验 a one-sided (or one-tailed) test
  - 备择假设中的参数值与零假设中的参数值的差异存在方向性
    - 更大? 更小? (而不是没有方向性的 “不一样” )
    - 例如:
      - $H_0$ : 男性与女性身高一样;
      - $H_A$ : 男性身高 > 女性身高;
  - 单侧检验中的备择假设及其方向性必须在获得数据前设定
    - 而不是依据数据结果
    - 数据本身不能用来预测偏离的方向
  - 当数据偏离的方向符合 $H_A$ 中描述的方向,  $H_0$ 被拒绝;

## 5. 课堂练习 Practice Problems



课堂测试-05



- Q1: 设想你正在使用随机数据样本来检验一个零假设。请回答以下说法正确与否：与抽样误差小的估计值相比，抽样误差大的参数估计值会导致检验的I类错误率更高。
  - 正确
  - 错误



## 5. 课堂练习

- 假设一个统计检验的零假设为真。请说出下列说法正确与否：
  - Q2: 如果你通过检验拒绝了这个  $H_0$ , 你就犯了第二类错误。
    - 正确
    - 错误
  - Q3: 如果你未能通过检验拒绝这个  $H_0$ , 你就犯了 I 类错误。
    - 正确
    - 错误





## 5. 课堂练习

- Q4: 把显著性水平 $\alpha$ 的值从 0.05 降到 0.01, 对以下各项有怎样的影响?
  - a. 犯第一类错误的概率?
  - b. 犯第二类错误的概率?
  - c. 检验的功效?
  - d. 样本大小?

(降低? 增大? 无影响? )

## 6. 小结 – 假设检验

- 假设检验是指将实际数据与特定零假设为真的期望结果进行比较的统计过程。如果数据与零假设成立时的预期数据相比过于异常，则拒绝零假设；
- 假设检验的步骤：
  - 1. 提出假设 ( $H_0$  &  $H_A$ )
  - 2. 利用数据计算检验统计量
  - 3. 确定 P 值
  - 4. 得出适当的结论
  - 5. 报告结果

## 6. 小结 – 假设检验

- 零假设 $H_0$ 是关于总体参数值的一种具体说法（假设为真）；一般来说，零效假设是人们倾向于拒绝的一个陈述；
- 备择假设 $H_A$ 包括零假设中所述值之外的所有其它可能的总体参数值，通常包括在生物学上比零假设中的值更有趣的可能性；
- 检验统计量是根据事实数据计算的一个数字，用于评估该数据与零假设的预期结果的匹配程度；
  - 零分布是在假定零假设为真的前提下，检验统计量结果的抽样分布；
- P 值是在零假设 $H_0$ 为真时获得观察数据（或更极端结果）的概率；
- 在双侧检验中，备择假设 $H_A$ 包括零假设 $H_0$ 所指定参数值两侧的参数值。

## 6. 小结 – 假设检验

- I 类错误是拒绝为真的零假设的错误，II 类错误是没有拒绝为假的零假设；
  - 显著性水平  $\alpha$  设定了犯一类错误的概率；
  - 生物学研究中通常将  $\alpha = 0.05$  视为可接受的错误率；
  - 功效 是拒绝为假的零假设的概率；通常增大样本量可以提高功效；
- 不能拒绝零假设通常不能断定假设是真还是假；
  - 因此建议在给出假设检验结果时，应尽可能附上相关参数的置信区间；
  - 置信区间则衡量效应的大小；
- 阅读资料：Whitlock & Schluter - Ch6