

## Biostatistics (BIOL0031132104) - Assignment #1

(released on Sept. 26<sup>th</sup>, due on Oct 7<sup>th</sup>)

1. In a study of diet preferences in leafcutter ants, a researcher presented 20 randomly chosen ant colonies with leaves from the two most common tree species in the surrounding forest. The leaves were placed in piles of 100, one pile for each tree species, close to colony entrances. Leaves were cut so that each was small enough to be carried by a single ant. After 24 hours, the researcher returned and counted the number of leaves remaining of the original 100 of each species. Some of the results are shown in the following table. Using these results, the researcher estimated the proportion of *Spondius mombin* leaves taken as 0.65 and concluded that the ants have a preference for leaves of this species. [40 marks]

在一项关于食叶蚁饮食偏好的研究中，研究人员向 20 个随机选择的蚁群提供了周围森林中两种最常见树种的树叶。树叶以 100 片为一堆，每种树叶一堆，堆放在蚁群入口附近。树叶被裁剪得足够小，以便一只蚂蚁搬运。24 小时后，研究人员返回计算每个树种 100 片树叶中剩余的数量。部分结果见下表。根据这些结果，研究人员估算出 *Spondius mombin* 的叶片比例为 0.65，并得出结论：蚂蚁偏爱这种叶片。

Tree species	Number of leaves removed
<i>Spondius mombin</i> (黄槟榔青)	1561
<i>Sapium thelocarpum</i> (乌柏属的某种)	851
Total	2412

- a. Identify the two variables whose association is displayed in the table. Which is the explanatory variable and which is the response variable? Are they numeric or categorical? 找出表格中显示其关联的两个变量。哪个是解释变量，哪个是响应变量？它们是数字变量还是分类变量？ [20 marks]

Tree species (树木物种) 是解释变量, number of leaves removed (被搬运走的叶片数量) 是响应变量; 其中树木物种是分类变量, 被搬运走的叶片数是数字变量。 [每个回答 5 分]

- b. Why do the 2412 leaves used in the calculation of the proportion not represent a random sample? 为什么计算比例时使用的 2412 片树叶不代表随机样本？ [10 marks]

随机样本需满足个体间相互独立且有相同概率被抽中的特征, 而 2412 片树叶是指两个树种被搬运走的叶片数的总和, 但该研究中的随机样本是指研究人员所放置的两个树的 20 个叶片堆 (针对 20 个蚁群), 每个叶片堆中的叶片相互并不独立, 计算比例时应该针对每个蚁群搬走的每个树种的叶片占该叶片堆总数 (100 片) 的比例, 获得 20 个蚁群针对每个树种搬运叶片数的比例。

[解释随机样本特征是什么占 3 分, 解释研究中的随机样本和 2412 片树叶区别占 7 分]

- c. Would treating the 2412 leaves as a random sample most likely affect the accuracy of the estimate of diet preference or the precision of the estimate? 将 2412 片树叶作为随机样本最有可能影响饮食偏好估计值的准确性或估计值的精确性吗? [10 marks]

该研究中 20 个蚁群是随机样本, 使用 2412 片作为随机样本会影响计算每个树木物种被每个蚁群搬运走的叶片数的比例, 因此会影响饮食偏好估计值的准确性。[回答影响准确性占 5 分, 解释原因占 5 分]

2. The following data are measurements of body mass, in grams, of finches captured in mist nets during a survey of various habitats in Kenya, East Africa (Schluter 1988). Read the data file and make calculations and a figure in R. [60 marks]

以下数据是在东非肯尼亚调查各种栖息地时用雾网捕获的雀类的体重测量值 (单位: 克) 。数据文件请见 'chap03q04KenyaFinches.csv', 请在 R 中完成计算和绘图。

Bird species	Body mass
Crimson-rumped waxbill (赤腰梅花雀)	8, 8, 8, 8, 8, 8, 8, 6, 7, 7, 7, 8, 8, 8, 7, 7, 7
Cutthroat finch (环喉雀)	16, 16, 16, 12, 16, 15, 15, 17, 15, 16, 15, 16
White-browed sparrow weaver (白眉织雀)	40, 43, 37, 38, 43, 33, 35, 37, 36, 42, 36, 36, 39, 37, 34, 41

- a. Calculate the mean body mass of each of these three finch species. Which species is largest, and which is smallest? 计算这三种雀类的平均体重。哪种最大, 哪种最小? [10 marks]

Crimson-rumped waxbill: 7.5 g

Cutthroat finch: 15.4 g

White-browed sparrow: 37.9 g

体重最大的物种是 White-browed sparrow 白眉织雀

体重最小的物种是 Crimson-rumped waxbill 赤腰梅花雀

[每个回答 2 分]

- b. Which species has the greatest standard deviation in body mass? Which has the least? 哪个物种的体重标准偏差最大？哪个最小？[10 marks]

体重标准差最大的物种是 White-browed sparrow 白眉织雀 (3.1 g)

体重标准差最小的物种是 Crimson-rumped waxbill 赤腰梅花雀 (0.6 g)

[每个回答 5 分, 若没有给出数值扣 1 分]

- c. Calculate the coefficient of variation (CV) in mass for each finch species. How different are the coefficients between the species? Compare the differences in CVs with the differences in standard deviation calculated in part (b). 计算每个雀类的体重变异系数 (CV)。各物种之间的系数有多大差异？将 CV 的差异与 (b) 部分计算的标准偏差的差异进行比较。[20 marks]

CV of Crimson-rumped waxbill: 8.29%

CV of Cutthroat finch: 8.04%

CV of White-browed sparrow: 8.19%

三个物种的 CV 的数值非常接近, 相比而言三个物种体重的标准差存在差别, 但计算 CV 是标准差除以了体重, 因此消除了体重大小对标准差的影响。[每个计算值 5 分, 比较差异 5 分]

- d. Choose one graph type to plot the data in the table to show the relationship between bird species and body mass. 选择一种图的类型绘制表格中的数据, 以表示雀类与体重之间的关系。[20 marks]

从 stripchart, boxplot, violin plot 中任选一个类型即可 [做出图 15 分, 各种图元素美观加分, 包括横纵坐标刻度/单位/标签等等]

