

DAD-IS系统中用于风险等级分类的 群体规模数据的收集和估计

——抽样方法

I. 介绍

家养动物多样性信息系统 (DAD-IS) 是支持动物遗传资源(AnGR)管理的全球性交流工具，用于动物遗传资源数据的信息交换。它为国家协调员提供了一个安全的手段，用以国家数据的输入、更新和访问。纵观 DAD-IS，许多品种的种群数据一直缺失。这些信息对于指导国家动物遗传资源管理和计算可持续发展目标指标 2.5.2(由粮农组织负责)至关重要。该指标通过品种的种群规模来判定品种灭绝的风险。截至 2021 年 9 月，在全世界范围内，不包含灭绝品种，61% 的地方品种被归类为未知风险状态，29% 处于风险级别，11% 不存在风险。数据差距的主要原因是缺乏国家一级的群体规模数据。开展全国性普查费用很高，因此与可持续发展目标的年度报告相比，不宜频繁开展。此外，普查通常不提供品种水平的信息。因此，各国需要具有成本效益的解决方案，使它们能够定期估计其农业生产率的群体规模。

本文件以具有成本效益的方式促进收集信息和估计种群规模而开发的，提供了简便方法和工具。这一方法已在涉及拉丁美洲和北非若干国家的两个试点项目中进行了测试。

在第一部分中，我们介绍了该方法的一般原理，包括需求、抽样和统计方法。第二部分着重于实践考虑，为该领域的实施提供参考。第三部分介绍为便利分析收集的数据和估计种群规模而开发的应用程序。

II. 方法的原理

1. 考虑因素和先决条件

在没有详尽的登记制度或大规模的牲畜普查的情况下，可以使用不同的方法来估计基于抽样的牲畜数量(FAO, 2011)。这些方法可分为以下几类：

- 根据走访的田间、农场或家庭的样本，直接计算动物的数量和类型。直接计数在时间、成本、容量要求和适用性方面都是非常苛刻的，这取决于现场的情况。
- 评估(问卷调查、家庭或小组访谈、主要信息提供者和农民组织/育种协会的访谈)基于相关利益攸关方在不同尺度上提供的估计。这些方法执行速度相对较快，成本低于普查，但这些方法面临的挑战在于难以量化、标准化和汇集数据。

这里提出的方法是基于各种假设和经验。首先，基于随机或半随机抽样的地方评估方法来估计种群规模，要比全面或大规模牲畜普查更具成本效益。

就取样而言，需要考虑到，就某一特定畜种，不同品种的农业生态和社会文化因素相关性在不同地区的分布是不均匀的。因此，有效的抽样方法需要考虑到这种多样性，同时也考虑到行政区域划分，以便切实考虑到在实地收集数据的问题。

该方法的原则是根据分层抽样原则，评估国内不同农业生态或行政地区之间品种的分布，以符合成本效益的方式获得可靠的种群数量估计。通过适当的分层，分层抽样可比相同规模的简单随机样本提供更精确的估计，从而提高成本效益。

在区域一级的估计可以基于直接计数或评价，最终考虑各种方法的组合。至于所选择的方法，必须考虑三个主要的先决条件：

- 首先，必须在国家和地区层面(如农业生态区或行政区域)以及再下一层级(如市级)提供相对较新的全国普查或其他可靠的物种水平估计数据。事实上，小规模的直接计数或评估预计不能提供国家一级的准确数字估计。然而，如果物种水平的种群大小已经知道，这些方法可以提供一个好的估计品种再分配，因此可以估计相应的种群大小的品种。
- 其次，应该已经对感兴趣的主要品种进行了描述，并就如何在实地区分这些品种达成了共识(粮农组织, 2012)。
- 第三，无论选择何种方法，都需要通过区域行业管理部门、兽医和推广服务机构或农民协会等机构，在阶层一级建立机构能力和提供支持。

如果上述三个条件中任何一个条件没有得到满足，则不能采用本抽样方法。

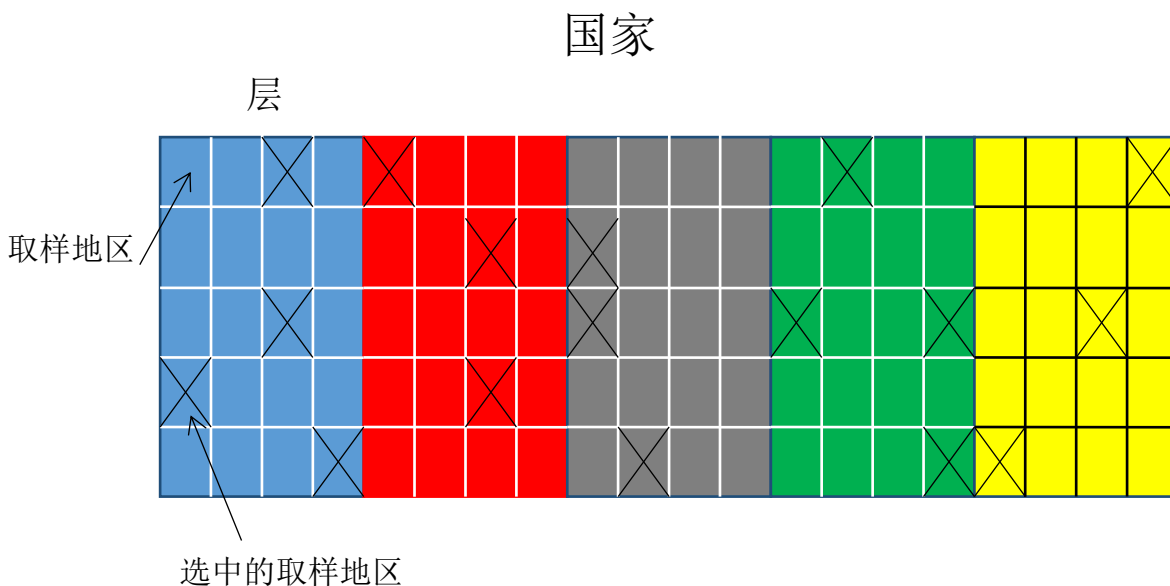


图 1: 同一国家的分层抽样计划

在每个层级中，抽样区域是以随机或半随机的方式选择的(图 1)，因此代表了整个阶层，而且调查人员也很容易地到实地访问。在每个抽样地区，品种种群的将根据机构合作伙伴提供的资料(例如当地饲养者或农民协会的评估、主要信息提供者提供的估计数字、屠宰场收集的资料等)或田间工作(例如直接点算、家庭访谈)得出。下一步，将利用从抽样地点得出的估计数来推算该层级的种群规模，并进一步推算全国的种群规模。

2. 应考虑的步骤

这种方法的实施需要几个连续的步骤:

第一步是制定策略，即确定数据收集过程是如何组织的，包括物种的选择覆盖，分层原则(农业生态区或行政单位)，用于评估品种在地层分布的方法(直接计数或评估)，合作伙伴参与。需要根据预算、时间、可用人力资源以及某些地区(如雨季)的可达性等方面的限制来讨论战略。这一步骤包括与国家和地方利益攸关方举行会议，以确保他们的参与，并最终培训实地工作人员以确定品种。

步骤 2 按照步骤 1 所设计的战略和方法，处理在不同地层和采样区实地收集信息的问题。

步骤 3 侧重于根据在每个采样点收集的信息估计国家一级品种种群规模。

在步骤 4 期间，应由牲畜和统计领域的专家对估计进行验证。根据步骤 3 和步骤 4 中确定的最终差异，可能需要进行补充工作以合并估计数。应当讨论经验教训和必要的调整，并将其记录在案，以改进今后使用的方法。

一旦验证，国家协调员应将数据输入到 DAD-IS。

3. 种群规模的统计估计

分层抽样是一种抽样方法，与简单随机抽样相比，抽样的精确度有所提高，从而减少抽样误差。

假设一个国家被细分为一个阶层。在全国范围内，一个给定物种的总种群大小估计为 N ，以 N_i 为第 i 层的种群大小，在第 i 层中，对 N_i 动物进行了抽样，其中一个比例的 p_{ik} 属于感兴趣的品种 k 。下面的公式(表 1)显示了如何估算品种的种群数量和置信区间。

表 1. 用于估计种群规模的公式

	公式
w_i : i 层的权重	$\frac{N_i}{N}$
\bar{p}_k : 品种 k 在国家一级估计的种群比例	$\sum_{i=1}^I p_{ik} \times w_i$
\bar{n}_k : 品种 k 群体规模估计数	$\sum_{i=1}^I p_{ik} \times N_i$
s_{ik}^2 : 品种 k 在层级 i 内估计比例的标准差	$(p_{ik} \times (1 - p_{ik}))^{\frac{1}{2}}$
S_{ek} : 比例估计 \bar{p}_k 标准差	$\sum_{i=1}^I w_i^2 \times \left(\frac{N_i - n_i}{N_i}\right) \times \left(\frac{s_{ik}^2}{n_i}\right)^{\frac{1}{2}}$
按照群体规模估计数 \bar{n}_k 划分的 95% 置信区间 上限	$(\bar{p}_k + 1.96 \times S_{ek}) \times N$
按照群体规模估计数 \bar{n}_k 划分的 95% 置信区间 下限	$(\bar{p}_k - 1.96 \times S_{ek}) \times N$

III. 实践中的执行

这一战略的设计和执​​行需要与该领域的畜禽品种专家以及统计专家进行初步讨论，因为这些决定将对种群规模估计和这些估计的准确性产生影响。数据需要从指定地层的代表性样本收集，以便充分估计品种分布。

1. 畜种的选择

该方法可以应用于选定的畜种，目的是估计某国该畜种所有品种的种群规模。畜种的选择将取决于其规模、社会经济重要性以及监测和管理能力。选择何种畜种进行调查，取决于该畜种的数量规模、特定的文化重要性、现有的管理(例如农民和品种协会的存在)和能力(可以动员的现场专家)。

畜种的选择决定了利益相关者参与实地考察工作。

2. 抽样目标：理论与实践

就动物数量而言，确定抽样规模取决于研究的目标，以及统计（环境的精度和结构）和实用性（环境约束，资金，时间和人力资源的可用性等）的考虑因素。（斯鲁斯菲尔德，1995年）。

表 2 显示了在理想情况下（随机抽样下的大量总体）根据样本数量估计品种频率的预期理论置信区间。预计该间隔将增加（因此精度预计会降低），品种的频率接近 50%。如果对于一个特定的畜种，只有少数品种，每个品种都有相对较高的频率 (>10%)，那么对几百只动物进行采样可以提供理论上它们在物种中的重要性的粗略概念。在实践中，由于实地抽样的非随机性（抽样的非随机性）以及正确识别品种的潜在困难，估计的精度预计会降低。因此，如果目标是定期监测品种种群，建议对该物种内至少 5 000 至 10 000 只动物进行取样。然而，即使对 10 000 只动物进行采样，预期频率仅为 1%的品种精度也将接近 0.2%，即预期频率的 20%。因此，可能需要根据先验知识对抽样进行调整，以更好地考虑小型本地品种（见以下各节）。

表 2.估计具有所需固定置信限（置信水平为 95%）的品种频率所需的样本量（斯鲁斯菲尔德，1995）

预期频率	所需的绝对精度		
	10%	5%	1%
10%	35	138	3457
20%	61	246	6157
30%	81	323	8067
40%	92	369	9220
50%	96	384	9604
60%	92	369	9220
70%	81	323	8067
80%	61	246	6157
90%	35	138	3457

3. 分层选择

理想情况下，分层应设计为与全国范围内不均匀品种分布的最佳匹配。由于经常描述本土品种适应特定的环境条件，品种分布模式可能在某种程度上与农业生态区相匹配。因此，在农业生态区划分国家可能是最佳的分层战略。然而，国内的行政区域也可以构成相关的分层单位，因为这些单位可能与社会环境差异相匹配，同时也与行政和治理单位相对应。如果抽样执行战略在很大程度上依赖于区域机构，那么使用行政单位而不是生态农业单位是有意义的。根据具体情况，其他因素，如生产系统或主要族裔群体的空间分布，最终可能会被考虑用于分层的定义（粮农组织，2011 年）。

表 3.基于两个理论场景的抽样工作重新分区

	畜种种群 规模总计	方案 A 抽样规模 (无假设)	方案 B 抽样规模 (层级 3 异质, 且有许多小品种)
层级 1	600 000	6000	4500
层级 2	200 000	2000	1500
层级 3	200 000	2000	4000

从理论上讲，抽样工作应与特定层级的种群重要性成正比。然而，如果预计某一地层的取样特别多或富含种群规模小的土著品种，则可特别重视在某一地层取样。表 3 提供了一个例子，其中决定对总共 10 000 只动物进行抽样。如果没有具体的假设，抽样工作将按比例对每个地层进行，但可以更强调先验信息，即已知特定地层是异质的，并且与其他地层相比，拥有许多小品种（参见采样框架部分）。

4. 估算方法的选择

粮农组织指南“调查和监测动物遗传资源”（粮农组织，2011年）对估算方法进行了广泛描述。制定旨在定期收集品种种群数据的有效战略需要决定使用的方法以及是否和如何最好地将它们结合起来。虽然评估通常更快，成本更低，耗时更少，但它们不那么精确，因为数字是从第三方推断出来的，例如，第三方可能对如何定义品种有不同的看法。进行评估的合作伙伴还应提供评估的规模，即进行品种重新划分的动物数量。在评估方法中，问卷调查可以以正式的方式增加参与的合作伙数量，从而增加样本量，但是它们的灵活性也较低，有时可能由于后勤和社会学问题（农民的识字程度……）而容易产生偏见。

根据当地情况，可以选择不同的方法并最终混合使用，因为对于特定国家，一种方法可能与一个阶层相关，而与另一个阶层无关。在同一地区本地实施两种不同的方法也可以是偶尔检查估计可靠性的一种方式（见方框 1）。在实践中，在南美和北非试点项目中，该战略主要依靠直接统计、住户调查和关键人评估。审定会议（见上文第 4 步）旨在通过补充评估方法合并实地作出的估计。

方框 1. 将抽样结果与哥伦比亚的其他估计数进行比较

在哥伦比亚的种群评估期间，可以将少数品种的抽样方法提供的结果与育种者协会和保护机构的估计数进行比较。在四个品种中的两个品种中，专家的估计值在抽样估计值的 95% 范围内，而对于一个品种（卡萨纳雷尼奥牛），两个品种之间的估计之间存在相当大的差异，因此需要进一步验证。

表 4. 基于抽样方法的种群估计，与其他估计值的比较

畜种	品种	抽样估计 (95%上限和下限)	其他 估计
牛	Casanareño	3100 (2873 - 3327)	977
牛	Romosinuano	3197 (2911 - 3483)	3709
牛	Sanmartinero	2360 (2162 - 2558)	2547
山羊	Criolla	9673 (8976 - 10370)	9440

5. 采样框架的实际考虑

根据农业生态特征、管理系统以及经济和社会文化因素，同时还考虑所使用的估计方法，样地地点可以由牛群、农户、村庄甚至小型行政单位组成。当然，直接计数方法更难以应用于大型样品位点。给定地层内的采样框架可以考虑一层、两层甚至三层，例如，第一层可以是市镇，第二层对应于市内的农场，作为采样区。

方框 2.厄瓜多尔的采样分层

为了评估厄瓜多尔的品种种群，采样框架基于与不同生态农业地区相连的三个区域，即科斯塔（太平洋沿海平原）、塞拉利昂（高地和山区）和东方（从安第斯山脉东部山坡到亚马逊雨林）。如图 2 所示为半随机抽样的抽样区域。例如，考虑到该地区的牲畜密度极低，与 Yasuni 国家公园和一些土著领土相对应，在奥连特地区东部的抽样受到限制。

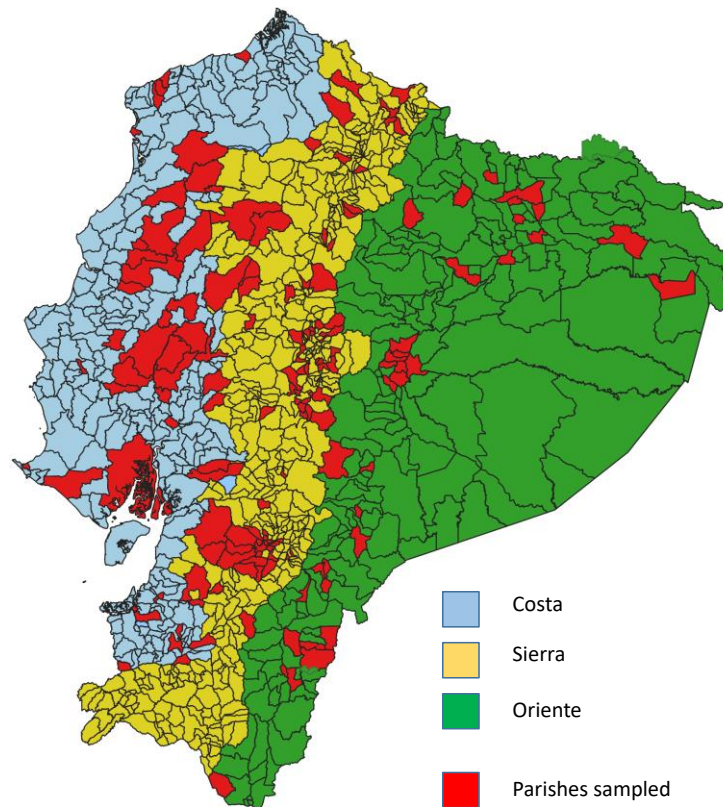


图 2.为评估厄瓜多尔三个生态农业地区的品种种群而抽样的区域

确定样本站点内的样本量是一项复杂的任务，必须考虑到（村庄之间品种重新划分的差异）和采样点内部的预期异质性（例如，特定村庄内接受访谈的家庭之间拥有的品种差异）之间的预期异质性。增加采样点的数量预计将减少与第一个组成部分相关的差异，而增加采样地点内计数的动物数量将减少第二个组成部分。

表 5.三个南美国家的采样框架和物种覆盖率

国家	层级	抽样单位	动物采样（种群覆盖率）
哥伦比亚	5 个农业生态区	25 Municipios (300 份调查问卷)	5908 鸡 (0.004%)
			43401 牛 (0.19%)
			430 水牛 (0.14%)
			3921 山羊 (0.34%)
			2834 马 (0.2%)
			3311 绵羊 (0.23%)
			1194 猪 (0.02%)
厄瓜多尔	3 个农业生态区	113 Parroquias (1100 份调查问卷)	197442 鸡 (0.34%)
			108347 牛 (2.51%)
			74764 豚鼠(1.48%)
			777 山羊 (1.96%)
			6000 马 (2.07%)
			3151 绵羊 (0.81%)
			43 骆驼 (0.11%)
巴拿马	5 个农业生态区	43 Corregimientos (7476 份调查问卷)	21828 牛(1.26%)
			469 山羊 goats (5.61%)
			553 绵羊 (2.96%)
			7648 猪 (2.37%)

表 5 说明了在巴拿马、哥伦比亚和厄瓜多尔采用的取样框架。物种数量和生态农业区因国家而异。对于这三个国家，小型行政单位（分别称为市政单位、帕罗基亚斯和科雷吉姆托斯）被定义为抽样单位，分别占哥伦比亚、厄瓜多尔和巴拿马抽样单位总数的 2%、8%和 6%。田间操作员的任务是使用直接访谈和实地问卷调查来访问这些采样点内的每个农场，种群覆盖率在实践中很少是详尽无遗的。此外，还与兽医和当地推广人员以及育种者协会进行了磋商，用这些记录完善了现场信息。这导致一些调查数量超过抽样单位的数量。根据物种的不同，采样的动物总数从 43 只（厄瓜多尔的骆驼科动物）到 197442（厄瓜多尔的鸡）不等。由于某些物种的抽样规模有限，这一过程没有得出所有品种和物种的种群估计数。例如，厄瓜多尔实际上在 DAD-IS 中报告了 8 个物种中只有 4 个品种的 36 个品种的种群估计数，其中大多数是牛。

6. 具体实例

该方法具有灵活性，允许根据具体情况调整方法或取样策略。例如，已知集中在有限地理区域的小型当地品种可能需要单独处理，以避免这些品种在一般抽样过程中被忽视。在拉丁美洲试点项目中，报告了在孤立领土上发现的这种种群规模较小的情况，其估计数是根据实地直接计算得出的，然后将这些数字直接列入国家总估计数。

利益相关者团体(如负责特定品种的育种组织或农民)能够提供其专业领域(品种或位置)的种群估计的精确数字。此外，一些品种可能有他们的种群的一部分非常好的监测，例如通过核心群，或商业性的品种登记或育种计划。相应的群体数据一般比较详尽，可以直接加入到群体估计中。

III. 估计品种种群规模的工具

为了便于估计种群大小，Red Conbiand 开发了一个计算机应用程序，以支持在拉丁美洲的试点项目中储存信息和估计种群大小。此 Access 数据库与 Windows 兼容，可以用三种语言访问。

需要输入以下类型的数据：

- 农业生态区的类型和采用的地区。
- 研究的动物品种及品种。
- 在国家一级及农业生态区一级每个畜种的普查
- 每个品种和采样区的动物数量。

抽样必须以分层的方式进行，应用程序的计算基于：

- 每个品种相对于物种的比例的估计：a) 在地区一级，b) 在国家一级。
- 计算该地区 and 该国估计种群规模的置信区间。

如对软件有需求，请写信至：DAD-IS@fao.org。

参考书目

Ayalew, W., Dorland, A. V., & Rowlands, G. J. 2004. Design, execution and analysis of the livestock breed survey in Oromiya Regional State, Ethiopia.

FAO. 2011. Surveying and monitoring of animal genetic resources. FAO Animal Production and Health Guidelines. No. 7. Rome.

FAO. 2012. Phenotypic characterization of animal genetic resources. FAO Animal Production and Health Guidelines. No. 11 Rome.

Rowlands, J., Nagda, S., Rege, E., Mhlanga, F., Dzama, K., Gandiya, F., Hamudikwanda, H., Makuza, S., Moyo, S., Matika, O., Nangomasha, E. & Sikosana, J. 2003. A report to FAO on the design, execution and analysis of livestock breed surveys – a case study in Zimbabwe. Nairobi, International Livestock Research Institute (available at <http://agtr.ilri.cgiar.org/sites/all/files/library/docs/FAOAndILRIZimbabewReport.pdf>).

Scholtz, M. M., Bester, J., Mamabolo, J. M., & Ramsay, K. A. 2008. Results of the national cattle survey undertaken in South Africa, with emphasis on beef. *Appl. Anim. Husb. Rural Dev*, 1, 1-9.

Thrusfield, M. 2005. *Veterinary epidemiology*. 2nd Edition, Blackwell Science, Oxford, 178-198.