值班主任：姚克成 责任编辑：刘英楠 新闻热线：(010)63376822 E-mail：zgxxbtjlt@163.com

抓实劳动力调查数据质量常态化自查

■ 杨沫

劳动力调查是监测就业形势、研判民生政策成效的基础性调查，数据质量是调查工作的生命线。以往质量自查多在上级抽查前夕集中开展，重资料补全、轻源头防控，容易出现“查一次、好一阵”的现象。笔者结合基层实操，就常态化推进数据质量自查谈几点思考。

机制嵌入日常，让风险巡检“常态化”

常态化自查，核心是把巡检固定下来、嵌进每月流程，而不是等抽查前突击补资料。实行“月自查、月复盘”：调查员每月完成样本户回访自查，辅助调查员对片区问卷开展交叉互审；市、县两级按季度抽取样本，远程核验与实地抽查相结合。前置标记风险户：重点关注流动人口、季节性务工、灵活就业、经营收入波动住户，标记为高风险样本，提高回访频次，台账动态更新。自查重点从“核对数字”转向“识别风险”，把问题化解在调查周期前端，避免年终集中整改、积重难返。

核户核人核状态，把住样本真实“第一关”

样本是劳动力调查的根基，户内人员信息失实是高频质量短板。自查首要抓好“三核”。核户，核查住宅是否真实存在，甄别空户、拆迁户、长期外出户，按规范启动换户流程，杜绝“虚拟户”。核人，逐户核对常住人口，重点排查户口在本地但长期不在此居住、多人合住、临时租住群体，厘清“居住界定”口径，防止漏记、多记。核状态，核查人员年龄、就学、退休身份，区分常住人口与短期访客，从源头规避人口基数错误。同步建立样本变动台账，每次自查更新居住变动、外出务工、户内增减记录，实现样本动态闭环管理。

交叉校验核心指标，深挖就业信息“逻辑矛盾”

核心指标集中在就业状态、工作时间、行业职业、未工作原因，单看一条数据往往看不出问题，必须强化指标间交叉校验。一看就业与工时是否匹配，对有工作但工时偏少、临时用工、零工、平台就业对象重点复核，准确区分就业、失业、非劳动力。二看行业职业填报是否对路，针对本地算力企业、物流、文旅季节性岗位等新业态用工，核实岗位描述，纠正行业职业错填。三看失业原因是否分清，区分主动辞职、被动失业、暂时停工，避免失业与非劳动力混淆。对逻辑异常问卷建立疑点清单，做到“一户一说明”，坚持入户回访核实原始情况，杜绝坐在办公室简单改数。

压实两级责任，做实问题溯源“闭环整改”

基层调查员直接接触住户，有可能出现理解偏差、代填估报、电话代访等问题，自查必须落到人头、追到源头。压实调查员自查责任，每份问卷填报完成先自我复核，模糊信息再次回访确认；推行片区交叉互查，不同调查员互抽对方样本，跳出自身惯性思维，更容易发现口径理解偏差。回访兼顾住户配合度，实地入户与电话回访相结合，随机抽查与定向疑点回访相结合。做实问题溯源，建立分类台账，将问题归入口径理解、入户时机、住户配合、样本管理等类型，定期复盘高频错误，形成风险提示清单，在月度业务会上通报，针对性开展小范围实操培训。自查留痕同步跟上，保存回访记录、佐证材料，做到可追溯、可复核，实现“发现一类问题、规范一类口径”。

（作者单位：国家统计局中卫调查队）

如何做好促销旺季采价

■ 朱施怡

“双十一”等电商大促及线下实体店打折旺季，满减优惠、预售定金、买一赠一、跨店凑单、直播专属价等促销模式叠加交织，价格形成机制复杂多变，给采价员准确判断真实成交价带来不小挑战。结合基层工作实践，笔者就如何规范采集促销期间价格数据谈几点体会。

以精准识变，把好促销规则“识别关”

促销花样繁多，把准活动的时间与门槛是剥离真实价格的第一步。采价中，首先严守时间门槛：实体店返券打折等活动持续时间超过7天，网络促销超过24小时，方可纳入采集范围；短于该时限的瞬时活动，直接沿用上一采价日价格。其次算清满减账：对于商场常用的“满额返券”，直接向营业员询问“若不返券最低能打几折”，按底价计算，若无法折现则记录原价；对于“满a元减b元”活动，先观察单件规格品是否满足门槛，满足则采减免后价格，不满足则采实际成交价。只有把这些基础规则吃透，才能避免把短期噱头当成常态价格。

以精细拆解，算清让利折算“明白账”

面对复杂的赠品与优惠形式，必须统一折算口径。针对“买一赠一”，若赠品与原规格品同品种，按容量或重量折算实际单价；若赠品非同一品种，则不考虑赠品价值，直接采集原规格品价格。若遇“预付定金享受优惠”，直接采集当天的销售价格，不采集“定金加尾款”折算出的虚拟低价。同时明确两条红线：对于仅面向特定消费者（如会员专属）的折扣不采，只取普通消费者均能购买的单件成交价；遇到商品临近过期或限量供应的打折，原则上采集原价，仅当打折属于普遍现象时才采折扣价。只有把折算标准统一、把口径定准，数据才能真实反映市场走势。

以规范操作，落实一线采集“执行链”

电商与实体店页面、价签更新快，采价员如果操作不统一，同一商品采出的价格可能相差甚远。对此，需要推行标准化的采集流程。对于服装鞋类等受促销影响剧烈的规格品，必须采集实际成交价；当发现规格品价格降至上市初始价格七折以下时，需警惕其是否因过度打折失去代表性，原则上应考虑更换规格品，避免价格越换越低。采价时遇到饮用水、香皂等连包整箱售卖的规格品，采集后必须换算成目录规定的基础单位价格。业务骨干对重点平台和商场开展跟采抽查，现场纠正操作偏差，确保一线采集严格符合制度规范。



图为工作人员在超市进行采价。

以闭环审核，筑牢数据质量“防火墙”

促销期间价格波动幅度大、变动频率高，审核环节必须跟上节奏、把住关口。需要建立“即报即审、当日清零”机制，对采价员上报的每一笔促销价格逐条复核。重点排查三类异常：一是价格降幅明显超出合理区间的，核实是否误采了限量秒杀等极端价格；二是同一商品在不同平台或网点价差过大的，核实是否混淆了会员价与普通成交价；三是促销结束后价格迅速回弹的，回溯采价时点是否准确、是否误将活动预热价当作正式活动价。对存疑数据要求补充截图佐证，及时联系采价员核实。每期采价结束后，将促销规则判断错误、单位未换算等典型问题整理成易错清单，在工作群通报提醒。通过初审、复审两级把关和异常回溯机制，确保促销季采集的每一笔价格数据都经得起推敲。

（作者单位：国家统计局沈阳调查队）

统计科普

PPI的编制方法（之一）

■ 王有捐 安冰洁

一、基本情况

PPI的编制由国家统计局组织实施。全国和各省（区、市）PPI由国家统计局编制并发布。2010年，PPI统计调查制度进行了改革，自2011年开始计算以2010年为基期的定基价格指数，并且每5年进行一次基期轮换。根据统计制度规定，2021—2025年编制和发布以2020年为基期的价格指数。

二、调查范围

PPI调查采取重点调查与典型调查相结合的调查方法，在全国31个省（区、市）中，抽选年主营业务收入2000万元及以上的工业企业作为调查对象。目前，全国有4万余家工业企业作为代表企业参与调查。

三、调查分类及主要产品、规格品目录

PPI调查分类及主要产品目录（简称调查目录）包含了41个工业行业大类（实际调查40个工业行业大类），207个工业行业中类，666个工业行业小类，1310个基本分类。PPI调查目录由国家统计局制定，各地可根据实际情况对产品、规格品进行补充。PPI调查目录编制的基本原则：

一是根据国民经济行业分类标准和《统计用产品分类目录》确定调查目录。

二是确保覆盖全部工业行业的主导产品，保证调查目录的覆盖面和代表性。

三是在使用上尽可能满足工业统计对分行业价格指数的需求。

四是粗细结合，对一些未来一段时期可能持续存在的重要产品，特别是重要生产资料，分类尽可能详细。

四、代表产品（规格品）的选择

编制PPI，是以一组代表产品（规格品）的价格变动来反映全部产品的价格变化趋势和变动幅度。选择代表产品（规格品）应遵循以下原则：

一是按工业行业选择基本分类和代表产品（规格品）。各个主要工业大类行业（其他采矿业除外）和重点中类行业，一般应选择足够数量的基本分类和代表产品（规格品），以使价格指数较好地反映各行业价格变动情况。

二是选择对国计民生影响大的产品。一般来看，销售产值大的产品对国计民生影响较大，应选择作为代表产品（规格品）。

三是选择生产较为稳定的产品。一旦被选为代表产品（规格品），就要连续调查较长时期。因此选择代表产品（规格品）时一定要考虑其生产的稳定性，试生产、经济寿命短的产品不应被选为代表产品（规格品）。

四是选择有发展前景的产品。部分产品如电子产品、生物制品、新材料等，尽管其当期销售产值较小，但随着时间的推移，其市场份额占比可能逐步升高，可选为代表产品（规格品）纳入调查。而部分产品尽管一时销售产值较大，但已是国家明令淘汰或是将被市场淘汰的，则不应选为代表产品（规格品）。

五是可以选择具有地方特色的产品。一些产品具有地方特色，也可以被选为代表产品（规格品）。

（摘自《领导干部应知应会主要统计指标诠释》，中共中央党校出版社、中国统计出版社）

统计学苑

在不确定的世界里寻找确定

——统计人生的回归、方差与推断

■ 朱利平



朱利平，中国人民大学教授、博士生导师，长期从事高维度和非线性大规模数据分析方法、理论和应用研究工作。

统计学是一门与不确定性相处的学问。它不承诺消除偶然，也不试图把世界简化为一条没有噪声的直线。它所做的，是在纷繁现象中寻找结构，在有限信息中形成判断，在变化与波动之间辨认那些较为稳定、可以依赖的部分。

也正因为如此，统计学不仅是一套处理数据的技术，更是一种认识世界的方法。多年学习、研究和讲授统计学之后，我越来越感到，许多看似抽象的统计概念，其实与人的成长经验暗暗相通。回归、方差与推断，既是统计学中的基本语言，也可以成为理解人生的三个关键词。

一、回归：从耀眼的极值回到真实的坐标

“回归”是统计学中最常见的词之一。我们借助回归模型研究变量之间的关系，也借助它从嘈杂的数据中辨认趋势。统计学史上的“回归”一词，原本就带有一种朴素而深刻的观察：极端状态往往难以永久维持，许多现象会在时间中向更为通常、更为稳定的水平靠近。

年轻时，人们往往不愿意谈回归。我们希望自己成为分布右端最醒目的点，希望每一次考试都有更高的分数，每一项工作都有更响亮的回报，每一个阶段都比前一个阶段更加耀眼。校园中的评价体系相对清晰：课程有成绩，论文有结果，竞赛有名次，毕业有去向。努力与反馈之间，似乎存在一条可被感知的函数关系。

走出校园以后，生活却不再服从如此整齐的设定。因变量不再由别人预先规定，解释变量也常常难以度量。事业、家庭、健康、兴趣、责任、机遇，各自进入人生模型，却没有统一的量纲。一个人究竟应当追求什么、愿意为什么长期投入、如何在公共责任与私人生活之间配置时间，都不能仅靠外部排名来回答。

这时我们才逐渐理解，回归并不是失败，更不是从优秀滑向平庸。它首先意味着回到真实的坐标。掌声会停下，光环会减弱，身份和头衔也会随着环境改变。真正能够留下来的，往往是一个人在无人注视时仍愿意认真做事的习惯，是面对复杂问题时仍坚持求真的耐心，是在日常责任中不轻易敷衍的品格。

统计中的回归，是从局部波动中寻找总体趋势；人生中的回归，则是穿过外界评价之后重新辨认自己的内在尺度。一个人的价值，并不由某个时点的极值决定，而更像是由一条长时间序列所刻画。偶尔领先并不等于长期卓越，阶段性的沉寂也不意味着趋势已经终止。真正重要的，是在足够长的时间尺度上，能否持续积累、不断校正，并将自己的工作变成推动整体进步的一点力量。

均值附近并不等于无所作为。社会的改进，往往不是依靠少数耀眼的异常值独自完成，而是依靠大量普通人在各自位置上提高工作的标准。一个教师认真讲好每一堂课，一个调查员忠实记录每一份数据，一个基层统计人员守住数据质量的底线，一个研究者耐心验证一个并不夺目的结论，他们未必站在聚光灯下，却共同推动着整体均值向上移动。

回归因此也是一种成熟：不再把“与众不同”当作唯一价值，不再把外部赞许当作唯一反馈，而是在真实、平常和长期之中建立稳定的自我。

二、方差：没有波动，便没有信息

如果说均值描述了一个总体的大致水平，那么方差揭示的，则是这个世界如何分散、如何不同、如何充满可能。

统计学之所以存在，正是因为现实不是完全齐一的。若所有观测值都相同，很多问题便无需研究。正是差异、偏离和波动，使数据携带信息，使比较成为可能，也使我们能够进一步追问：差异来自何处，是随机扰动，还是结构性因素？是暂时现象，还是长期分化？

人生也有方差。有人很早找到方向，有人需要经过多次转弯；有人在稳定环境中积累，有人在不确定中开辟道路；有人因一次偶然机会改变轨迹，也有人在看似没有变化的岁月里完成缓慢而坚实的生长。真实的人生从来不是均值的简单展开，而是一条包含峰谷、残差与结构突变的曲线。

我们常常本能地厌恶方差。家长希望孩子少走弯路，年轻人希望职业选择一次成功，组织希望计划得到精确执行，个人希望努力能够迅速兑换为确定结果。然而，方差并不只是风险，它也是可能性的来源。没有差异，便无从识别；没有变动，便难以学习；没有一定程度的不确定，许多新的道路也不会出现。

当然，赞美方差并不意味着浪漫化风险，更不意味着把所有挫折都解释为成长。统计思维的价值，恰恰在于区分不同性质的波动。有些波动是可以通过制度设计、专业训练和风险管理加以降低的，例如调查误差、测量偏差、数据造假和决策中的盲目性；有些不确定性则是开放系统的内在属性，不能被彻底消除，只能被识别、估计和管理。

对个人而言，也需要作出类似区分。我们应当努力减少由懒惰、轻率和失信造成的“人为方差”，却不必把人生中一切偏离预期都看作错误。一次职业调整、一个未被采用的研究设想、一段暂时没有回报的积累，都可能在更长的时间尺度上显示出新的意义。

统计学家面对离群点，不会立即将其删除，而会先判断它究竟是记录错误，还是一个值得解释的新现象。面对人生中的“离群时刻”，我们也需要这样的耐心。有些异常确实需要修正，有些异常却可能提示我们：原有模型不够充分，既定道路并非唯一选择，对自我的理解也需要更新。

理解方差，意味着承认人与人之间存在真实差异，也意味着放弃用单一标准裁剪所有人生。一个健康的社会，不应只追求均值的提高，也应当容纳合理的差异，为不同才能、不同节奏和不同选择留下空间。统计的眼光使我们看到总体，也提醒我们不要让总体遮蔽个体；它要求我们寻找规律，也要求我们尊重每一个观测值背后的具体处境。（未完待续）

（作者系中国人民大学统计与大数据研究院院长）