

团体标准

T/ CSSC 2022—XXX

12345 政务服务便民热线 第三方评估通则

Third party evaluation general rules for 12345 government service
convenience hotline

2022-XX-XX 发布

2022-XX-XX 实施

中国商业统计学会 发布



版权保护文件

版权所有归属于该标准的发布机构。除非有其他规定，否则未经许可，此发行物及其章节不得以其他形式或任何手段进行复制、再版或使用，包括电子版，影印件，或发布在互联网及内部网络等。使用许可可于发布机构获取。

-

目 次

前 言 IV

引 言 V

1 范围 6

2 规范性引用文件 6

3 术语和定义 6

4 评估原则 7

5 评估指标体系 7

6 评估方法 9

7 评估程序 9

附录 A （资料性附录） 评估指标测量示例..... 11

附录 B （资料性附录） 客观赋权法..... 14

附录 C （资料性附录） 主观赋权法..... 17

附录 D （资料性附录） 数据采集方式..... 20

前 言

本文件按照GB/T 1.1-2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由XX提出。

本文件由XX归口。

本文件起草单位：XXX。

本文件主要起草人：XXX。

引 言

12345 政务服务便民热线（简称 12345 热线）是畅通社情民意表达渠道重要载体，在加强和创新社会治理、打造服务型政府过程中发挥着重要作用。

历经 30 多年的发展，越来越多的地方政府以 12345 热线为基础，整合多条热线，为群众提供“一号对外”服务。截至 2020 年 6 月，全国地级及以上城市基本上已全部设立 12345 政务服务便民热线，各区、县乃至街镇都设立了相关机构。

基于第三方视角的 12345 热线评估标准，帮助规范、统一各地对 12345 热线运行质量的评估，衡量当前 12345 热线的工作现状，也有利于对 12345 热线进行管理和指导，对推动 12345 热线发展、推进治理现代化具有重要意义。

12345 政务服务便民热线第三方评估通则

1 范围

本文件确立了12345政务服务便民热线第三方评估的原则，给出了评估指标体系、评估方法和评估程序。

本文件适用于12345政务服务便民热线第三方评估活动。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 33358—2016 政府热线服务规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

12345 政务服务便民热线 “12345” government service convenience hotline

12345 热线 12345 hotline

由政府及其相关职能部门设立的公共服务平台。

3.2

12345 热线服务 12345 hotline service

由12345热线（3.1）为组织或个人提供的咨询、诉求处置与办理等公共服务。

注：12345热线服务渠道通常包括语音呼叫号码为“12345”的电话，微信、微博等手机APP，以及网站、信箱等。

3.3

12345 热线服务提供者 12345 hotline service provider

12345热线（3.1）的运行管理机构。

3.4

机构 body

有特定任务和组成的法定实体或行政实体。

注：机构如组织、权力机关、公司和基金会等。

[来源:GB/T 20000.1—2014, 6.1, 有修改]

3.5

服务对象 service objects

向12345热线（3.1）咨询信息、反映诉求、提出意见建议的自然人、法人或其他组织。

[来源:GB/T 33358—2016, 3.4, 有修改]

3.6

服务人员 service staff

为服务对象（3.4）直接提供服务的12345热线工作人员。

3.7

第三方 third party

与12345热线服务提供者（3.3）没有行政隶属关系的法人或其他组织。

4 评估原则

开展12345热线第三方评估，应在确保评估的过程和结果客观、公正的基础上，进一步遵循以下原则：

- 1) 独立性原则：评估过程和结果应不受12345热线服务提供者干扰。
- 2) 代表性原则：评估数据采集应能反映12345热线服务的整体工作状态。
- 3) 一致性原则：评估数据处理应有统一的规则，保证评估结果规范一致。
- 4) 可追溯性原则：评估数据、评估过程应有记录，保证评估结果可追溯。

5 评估指标体系

5.1 构建原则

5.1.1 可操作性

评估指标应易于理解，便于数据的采集和处理。

5.1.2 典型性

评估指标应反映12345热线的整体状况，以最少的指标反映12345热线的服务状况。

5.1.3 可比性

评估指标应有较为广泛的适用性，便于不同层级12345热线之间的横向比较。

5.2 指标体系框架

12345热线第三方评估包括服务响应、服务规范、问题解决、服务创新4个一级指标，每个1级指标下设2个二级指标，具体指标体系构成见图1。

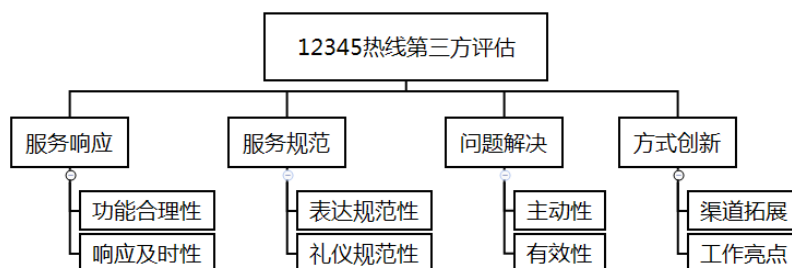


图1 12345热线第三方评估指标体系

5.3 指标说明与示例

12345热线第三方评估指标体系的一级指标、二级指标和指标说明，以及三级指标示例见表1。

表 1 12345热线第三方评估指标说明

一级指标	一级指标说明	二级指标	二级指标说明	三级指标示例
服务响应	12345 热线响应服务对象需求的情况。	功能合理性	服务功能设置的合理程度。	全天候人工服务提供
				交互应答设置 注：交互应答设置不宜超过 2 级，各轮选项设置不超过 5 个
		响应及时性	响应服务需求的及时程度。	接通率
				语音播报时长 转人工用时
服务规范	12345 热线提供服务时的规范程度。	表达规范性	服务人员话语表达的规范程度。	提供普通话服务
				语速适中，清晰报读
				表达流畅
				不随意插话，耐心倾听
		礼仪规范性	服务人员话语礼仪的规范程度。	自我介绍
				礼貌问好、称呼对方
				查资料前后礼貌致歉
				再次询问需求
				礼貌告别
				致谢
问题解决	12345 热线解决服务对象来电诉求的情况。	主动性	服务人员掌握、解决服务对象诉求的主动程度。	首呼解答（针对相同咨询场景）
				态度亲和
				情绪安抚
				询问细节
		有效性	服务人员解答，以及解决服务对象诉求的准确和有效程度。	有解答
				解答准确
				答复一致
				进度自助查询
				服务效能“好差评”
服务创新	12345 热线为更全面收集、高效解决服务对象诉求、提供更好的服务所采取的新措施、新方式。	渠道拓展	12345 热线服务提供者在原有电话端基础上进行的服务拓展及服务提供情况。	在线交流（通过手机 APP，如微信、微博等，以及门户网站统一提供在线交流服务）
				功能齐全（包括咨询、求助、投诉、举报、意见建议等功能）
				咨询有效（在拓展渠道内的咨询回复及时，解答有效）
				知识库自助查询（提供知识库，方便检索，内容更新及时）
				话务分析（能够对话务量进行基础性统计分析，并在相关热线渠道展示）
		服务亮点	12345 热线服务提供者自身的亮点服务功能。此部分可设置得分上限。	线上自助下单
				电话端、线上智能客服
				线上人工客服
				为企业服务
				外语服务
				特殊人群服务
				其他服务亮点

6 评估方法

6.1 评估模型

评估结果按式（1）计算：

$$S = \sum_{i=1}^4 X_i Q_i \dots\dots\dots (1)$$

式中：

S ——评估结果得分， $0 \leq S \leq 100$ ；

X_i ——第*i*个一级指标的评估得分；

Q_i ——第*i*个一级指标的评估权重。

X_i 按式（2）计算：

$$X_i = \sum_{j=1}^n Y_{ij} \cdot Q_{ij} \dots\dots\dots (2)$$

式中：

Y_{ij} ——第*i*个一级指标的第*j*个二级指标的评估得分；

Q_{ij} ——第*i*个一级指标的第*j*个二级指标的评估权重。

Y_{ij} 按式（3）计算：

$$Y_{ij} = \sum_{k=1}^m Z_{ijk} \cdot Q_{ijk} \dots\dots\dots (3)$$

式中：

Z_{ijk} ——第*i*个一级指标的第*j*个二级指标的第*k*个三级指标的评估得分；

Q_{ijk} ——第*i*个一级指标的第*j*个二级指标的第*k*个三级指标的评估权重。

各评估指标的测量方法参见附录A。

6.2 评估权重

各评估权重的取值范围应在0~1之间，同一层级的评估权重之和应为1。

根据不同的评估目的选择客观赋权法（参见附录B）或主观赋权法（参见附录C）确定评估权重。

7 评估程序

7.1 确定评估目的

不同的评估目的，将得到不同的评估结果。在实施评估前，应明确评估要达到的目的和要解决的实际问题。评估目的包括但不限于：

- 1) 发现12345热线提供的服务与服务对象诉求之间的差距；
- 2) 证针对12345热线改进措施的有效性；

-

3) 不同12345热线间的横向比较,为提出针对性的改进措施提供依据。

7.2 制定评估方案

评估方案应包括但不限于评估目的、评估指标、评估权重,以及评估数据的采集方式(参见附录D)、样本量、采集时间、采集人员。

7.3 采集评估数据

按照评估方案采集评估数据信息。

7.4 保存评估数据

采集到的原始评估数据应至少保存24个月,保障评估结果可追溯。

7.5 分析评估数据

对采集到的原始评估数据进行清洗,剔除问题数据,确保评估数据的准确合理。

7.6 给出评估结果

按第6章评估方法计算出评估结果。

7.7 撰写评估报告

对各级评估指标的得分进行汇总分析,撰写12345热线第三方评估报告。评估报告应包括但不限于以下部分:

- 1) 整体评估结论;
- 2) 成效与问题分析;
- 3) 问题改进意见;
- 4) 细分维度的研究分析成果,例如县市区、单位、部门等的分析结果。

附 录 A
(资料性)
评估指标测量示例

12345热线第三方评估指标测量示例见表A.1。

表A.1 12345热线第三方评估指标测量示例

一级指标	二级指标	采集方式	计算方法	评估结果
服务响应	功能合理性	神秘监测 问卷调查	神秘监测： $x = \begin{cases} 100 & A \geq N \\ \left(\frac{A}{N}\right) * 100 & A < N \end{cases}$ A:功能合理项符合规范的项目数 N:评估的项目数 问卷调查（主要针对热线功能的合理程度评估）： $x = \frac{100}{n} \sum_{i=1}^n \frac{a_i - 1}{4}$ N:评估的人数 a_i :取值1, 2, 3, 4, 5 1:非常不满意 2:比较不满意 3:一般 4:比较满意 5:非常满意	神秘监测： $0 \leq x \leq 100$ 问卷调查： $0 \leq x \leq 100$
	响应及时性	神秘监测 问卷调查	神秘监测： $x = \begin{cases} 100 & A \geq N \\ \left(\frac{A}{N}\right) * 100 & A < N \end{cases}$ A:响应及时项符合规范的项目数 N:评估的项目数 问卷调查（主要针对12345热线响应的及时程度评估）： $x = \frac{100}{n} \sum_{i=1}^n \frac{a_i - 1}{4}$ N:评估的人数 a_i :取值1, 2, 3, 4, 5 1:非常不满意 2:比较不满意 3:一般 4:比较满意 5:非常满意	神秘监测： $0 \leq x \leq 100$ 问卷调查： $0 \leq x \leq 100$

一级指标	二级指标	采集方式	计算方法	评估结果
服务规范	表达规范性	神秘监测 问卷调查	神秘监测： $x = \begin{cases} 100 & A \geq N \\ \left(\frac{A}{N}\right) * 100 & A < N \end{cases}$ A:表达规范项符合规范的项目数 N:评估的项目数 问卷调查（主要针对热线话务员在对话过程中表达规范程度评估）： $x = \frac{100}{n} \sum_{i=1}^n \frac{a_i - 1}{4}$ N:评估的人数 a_i :取值1, 2, 3, 4, 5 1:非常不满意 2:比较不满意 3:一般 4:比较满意 5:非常满意	神秘监测： $0 \leq x \leq 100$ 问卷调查： $0 \leq x \leq 100$
	礼仪规范性	神秘监测 问卷调查	神秘监测： $x = \begin{cases} 100 & A \geq N \\ \left(\frac{A}{N}\right) * 100 & A < N \end{cases}$ A:礼仪规范项符合规范的项目数 N:评估的项目数 问卷调查（主要针对话务员解答过程中话务礼仪的规范程度评估）： $x = \frac{100}{n} \sum_{i=1}^n \frac{a_i - 1}{4}$ N:评估的人数 a_i :取值1, 2, 3, 4, 5 1:非常不满意 2:比较不满意 3:一般 4:比较满意 5:非常满意	神秘监测： $0 \leq x \leq 100$ 问卷调查： $0 \leq x \leq 100$
问题解决	主动性	神秘监测 问卷调查	神秘监测： $x = \begin{cases} 100 & A \geq N \\ \left(\frac{A}{N}\right) * 100 & A < N \end{cases}$ A:主动性方面符合的项目数 N:评估的项目数 问卷调查（主要针对热线解决问题时服务的主动性程度评估）： $x = \frac{100}{n} \sum_{i=1}^n \frac{a_i - 1}{4}$ N:评估的人数	神秘监测： $0 \leq x \leq 100$ 问卷调查： $0 \leq x \leq 100$

一级指标	二级指标	采集方式	计算方法	评估结果
			a_i :取值1, 2, 3, 4, 5 1:非常不满意 2:比较不满意 3:一般 4:比较满意 5:非常满意	
	有效性	神秘监测 问卷调查	神秘监测: $x = \begin{cases} 100 & A \geq N \\ \left(\frac{A}{N}\right) * 100 & A < N \end{cases}$ A:有效性方面符合的项目数 N:评估的项目数 问卷调查（主要针对热线解决问题时服务的准确彻底程度评估）： $x = \frac{100}{n} \sum_{i=1}^n \frac{a_i - 1}{4}$ N:评估的人数 a_i :取值1, 2, 3, 4, 5 1:非常不满意 2:比较不满意 3:一般 4:比较满意 5:非常满意	神秘监测: $0 \leq x \leq 100$ 问卷调查: $0 \leq x \leq 100$
服务创新	渠道拓展	神秘监测 专家考评	神秘监测: $x = \begin{cases} 100 & A \geq N \\ \left(\frac{A}{N}\right) * 100 & A < N \end{cases}$ A:多样畅通方面符合的项目数 N:评估的项目数	神秘监测: $0 \leq x \leq 100$
	工作亮点	神秘监测 专家考评	神秘监测: $x = \begin{cases} 100 & A \geq N \\ \left(\frac{A}{N}\right) * 100 & A < N \end{cases}$ A:工作亮点方面符合的项目数 N:评估的项目数	神秘监测: $0 \leq x \leq 100$

附录 B (资料性) 客观赋权法

B.1 结构方程法

B.1.1 概念

结构方程是一种建立、估计和检验因果关系的模型。模型中既包含有可观测的显变量，也可能包含无法直接观测的隐变量。结构方程可以代替多种分析方法，分析单项指标对总体的作用和单项指标间的相互关系。

B.1.2 基本步骤

结构方程法的操作步骤：

- a) 确定适合的路径模型，作为研究初始的基准模型；
- b) 确定所建立的模型符合假设条件，利用Amos软件构件结构方程模型，并进行参数估计；
- c) 通过分析对模型进行拟合检验，如若通过，对模型进行修正，通过模型的路径系数确定指标权重。

B.1.3 优点

结构方程法的优点主要体现在以下几个方面：

- a) 隐变量可以通过一个或多个显变量进行测量；
- b) 根据可观测的变量可以计算出不可观测的隐变量之间的相互关系；
- c) 结构方程模型允许自变量和因变量含测量误差；
- d) 结构方程模型可以同时估计隐变量和显变量的关系及隐变量之间的关系；
- e) 结构方程模型既可以用图形也可以用数学模型进行描述。

B.1.4 局限性

结构方程法在应用过程中主要有以下几方面的局限性：

- a) 结构方程法对模型的接受没有统一标准，所以在某些情况下，很难拒绝某些模型，给模型选择带来困难；
- b) 影响结构方程模型的解释能力的主要问题是指定误差，但目前仍不能对其进行检验；
- c) 结构方程法对样本容量的要求较高，也要求模型必须满足识别条件，并且它不能处理真正的分类变量。

B.2 熵权法

B.2.1 概念

在信息论中，熵是对不确定性的一种度量。信息量越大，不确定性就越小，熵也就越小。在多指标决策问题中，某项指标的变异程度越大，信息熵越小，该指标提供的信息量就越大，在方案评估中所取得的作用就越大，该指标的权重也就越大。

B. 2. 2基本步骤

熵权法的操作步骤：

a) 将各个指标的数值进行标准化处理：假设给定了 k 个指标 $X_1, X_2, \dots, X_k$ ，其中 $X_i = \{x_1, x_2, \dots, x_n\}$ 。假设对各指标数据标准化后的值为 $Y_1, Y_2, \dots, Y_k$ ，那么 $Y_{ij} = \frac{x_{ij} - \min(x_i)}{\max(x_i) - \min(x_i)}$ 。

b) 求各指标的信息熵：根据信息论中信息熵的定义，一组数据的信息熵 $E_j = -\ln(n)^{-1} \sum_{i=1}^n p_{ij} \ln p_{ij}$ 。

其中 $p_{ij} = Y_{ij} / \sum_{i=1}^n Y_{ij}$ ，如果 $p_{ij} = 0$ ，则定义 $\lim_{p_{ij} \rightarrow 0} p_{ij} \ln p_{ij} = 0$ 。

c) 确定各指标权重：根据信息熵的计算公式，计算出各个指标的信息熵为 $E_1, E_2, \dots, E_k$ 。通过信息熵计算各指标的权重： $W_i = \frac{1 - E_i}{k - \sum E_i} (i = 1, 2, \dots, k)$ 。

B. 2. 3优点

熵权法的优点主要体现在以下几个方面：

- a) 客观性：相对主观赋值法，精度较高客观性更强，能够更好的解释所得到的结果；
- b) 适应性：可以用于任何需要确定权重的过程，也可以结合一些方法共同使用。

B. 2. 4局限性

熵权法在应用过程中主要有以下几方面的局限性：

- a) 缺乏各指标之间的横向比较；
- b) 各指标的权重随着样本的变化而变化，权数依赖于样本，在应用上限制。

B. 3 主成分分析法

B. 3. 1概念

主成分分析法是用较少的变量去解释原来资料中的大部分变量，将许多相关性很高的变量转化成彼此相互独立或不相关的变量。通常是选出比原始变量个数少，能解释大部分资料中变量的几个新变量，即所谓主成分，并用以解释资料的综合性指标。

B. 3. 2基本步骤

-

主成分分析法的操作步骤:

- a) 对原始数据进行标准化
- b) 确定主成分个数
- c) 解释主成分含义
- d) 用主成分及其方差贡献率构造主成分综合评估函数
- e) 计算综合评估函数的样本值并给出样本的排序

B. 3. 3优点

主成分分析法的优点主要体现在以下几个方面:

- a) 可消除评估指标之间的相关影响。
- b) 可减少指标选择的工作量。主成分分析法可以消除评估指标间的相关影响,所以在指标选择上相对容易些。
- c) 主成分分析中各主成分是按方差大小依次排列顺序的,在分析问题时,可以舍弃一部分主成分,只取前面方差较大的几个主成分来代表原变量,从而减少了计算工作量。

B. 3. 4局限性

主成分分析法在应用过程中主要有以下几方面的局限性:

- a) 在主成分分析中,我们首先应保证所提取的前几个主成分的累计贡献率达到一个较高的水平,其次对这些被提取的主成分必须都能够给出符合实际背景和意义的解释,否则主成分将空有信息量而无实际含义。
- b) 主成分的解释其含义一般多少带有点模糊性,不像原始变量的含义那么清楚、确切,这是变量降维过程中不得不付出的代价。因此,提取的主成分个数 m 通常应明显小于原始变量个数 p (除非 p 本身较小)。
- c) 当主成分的因子负荷的符号有正有负时,综合评估函数意义就不明确。

附录 C (资料性) 主观赋权法

C.1 德尔菲法

C.1.1 概念

德尔菲法是由评估机构选定有关专家，采用函询的方式，收集专家意见，加以综合整理，再将整理结果匿名反馈给专家。这样经过几次的征询及反馈，使各种不同意见逐步趋向一致，从而得出比较正确的结果。

C.1.2 基本步骤

德尔菲法操作步骤：

- a) 明确评估目标，明确进行效能评估的目标；
- b) 选聘专家：专家的构成要科学合理，应选择在标准制定、标准执行、标准研究等方面有独到见解的专家；
- c) 发布问题：发布需要专家评估的问题，分几轮进行评估，直到达到预期的收敛效果；
- d) 专家对问题进行评估。采用匿名评估，专家根据评估规则回答问题，并说明回答问题的依据，按照该程序完成对所有问题的回答。
- e) 对获取的专家知识进行处理：以专家的原始意见为基础，建立专家意见集成的优化模型，综合考虑一致性和协调性因素，同时满足整体意见收敛性的要求，找到群体决策的最优解或满意解，获得具有可信度指标的结论，达到专家意见集成的目的。

C.1.3 优点

德尔菲法的优点主要体现在以下几个方面：

- a) 匿名性：专家互不见面，可以消除专家之间的心理影响，做到充分自由地发表意见；
- b) 反馈性：德尔菲法要经过若干次的循环才能完成，各轮循环都是在精心控制下得到反馈；
- c) 收敛性：通过书面讨论，言之有理的意见会逐渐为大多数专家所接受，群体的见解会逐渐集中，呈现收敛的趋势。

C.1.4 局限性

德尔菲法在应用过程中主要有以下几方面的局限性：

- a) 从参与评估的专家来看，难以最大限度地发挥各自优势；
- b) 从评估的组织者来看，虽然处于主动地位，但工作量较大；

c)从专家和评估组织者的协调关系来看，组织者可以选择不同的专家组来完成不同环节的德尔菲过程，环节任务的不同和专家组性质的差异使得组织者和专家之间需要更多的时间和静力进行沟通和协调。

d)从整个评估过程来看，由于只涉及到部分环节，专家的判断意见涉及面相对较小，其贡献为单环节应用小于多环节应用，强约束性应用小于弱约束性应用。

C.2 层次分析法

C.2.1 概念

层次分析法又称 AHP 构权法 (Analytic hierarchy process, 简称为 AHP)，是将复杂的评估对象排列为一个有序的递阶层次结构的整体，然后在各个评估项目之间进行两两的比较、判断，计算各个评估项目的相对重要性系数，即权重。

C.2.2 基本步骤

层次分析法操作步骤：

- a) 构造判断矩阵
- b) 计算各个指标权重
- c) 进行一致性检验

C.2.3 优点

层次分析法的优点主要体现在以下几个方面：

a) 层次分析法中每一层的权重设置最后都会直接或间接影响到结果，而且在每个层次中的每个因素对结果的影响程度都是量化的，非常清晰、明确。这种方法尤其可用于对无结构特性的系统评估以及多目标、多准则、多时期等的系统评估。

b) 把定性方法与定量方法有机地结合起来，使复杂的系统分解，能将人们的思维过程数学化、系统化，便于人们接受，且能把多目标、多准则又难以全部量化处理的决策问题化为多层次单目标问题。

c) 层次分析法更讲求定性的分析和判断，所需定量数据信息较少。

C.2.4 局限性

层次分析法在应用过程中主要有以下几方面的局限性：

- a) 层次分析法只能从原有方案中进行选取，而不能为决策者提供解决问题的新方案。
- b) 定量数据较少，定性成分多，不易令人信服。
- c) 指标过多时数据统计量大，且权重难以确定。

-

d)特征值的特征向量的精准求法比较复杂。在二阶、三阶的时候，我们还比较容易处理，但随着指标的增加，阶数也随之增加，在计算上也变得越来越困难。

附录 D
(资料性)
数据采集方式

D.1 神秘监测

由经过严格培训的调查员，在规定或指定的时间里扮演成服务对象，对事先设计的一系列问题逐一进行评估或评定。由于需要被评定的服务人员，事先无法识别或确认“神秘顾客”的身份，故该调查方式能真实、准确地反映客观存在的实际问题。

D.2 问卷调查

通过设计者统一设计一系列问卷问题，收集服务对象对于某个12345热线的服务感受、态度行为特征价值观点或信念等信息。

问卷调查中常用抽样精度确定样本量，在95%的置信水平下，当样本量越大时，抽样误差会越小。

$$\Delta = \sqrt{\frac{t^2 PQ}{n}}$$

- n 样本规模
- t 概率度（一般置信度 95% 时，t=1.96，如置信度为 99%，t=2.6）
- Δ 极限统计误差
- P 概率，Q=（1-P）

$$\Delta = \sqrt{\frac{t^2 pq}{n}} = \sqrt{\frac{1.96^2 p(1-p)}{n}} \approx \sqrt{\frac{4 \times 0.25}{n}} = \frac{1}{\sqrt{n}}$$

图3 样本量计算公式

根据上述公式，不同精度要求下，样本量大小不同，详见下表：

表 D.1 不同精度要求下样本量大小表

精度（1-Δ）	极限误差（Δ）	样本量大小（n）
98.5%	1.5%	4500
98%	2%	2500
97.6%	2.4%	1800
95%	5%	400
94.2%	5.8%	300
92.9%	7.1%	200
92%	8%	156
90%	10%	100

-

D.3 专家考评

成立由公共管理、政务服务等方面专家组成的12345热线评估小组，根据指标体系对被评估的12345热线服务情况进行跟踪评估。