Глава 10

СТАТИСТИЧЕСКО ОЦЕНЯВАНЕ: ОТ ИЗВАДКА КЪМ ПОПУЛАЦИЯ

1. Защо е необходимо да изучаваме извадки?

Проучванията в медицината и здравеопазването обикновено имат за цел да оценят типичното ниво на някои показатели или честотата на разпространение на дадено явление сред определена популация. При проучване на масови явления доста често се оказва невъзможно да бъдат обхванати всички членове на съответната популация. Затова по необходимост изследователите често са принудени да наблюдават извадки и на тази основа да оценяват средни нива или честота на съответни явления в популация.

Основните причини за проучване на извадки се свеждат най-често до*:*

* ограничени финансови, времеви, технологични и други ресурси;
* липса на достъп до цялата популация;
* наблюдението на извадки може да е единствен възможен метод за набиране на информация.

Извличането на заключения за характеристиките на популацията от данните при наблюдение на извадка се нарича ***статистическо заключение*** *или* ***генерализация*** (обобщаване).

Статистическите заключения имат две основни задачи:

* *оценка на резултатите от наблюдение на извадка и извличане на изводи за параметрите на популацията;*

* *сравняване на резултати от наблюдение на две или повече извадки, т.е. проверка (тестуване) на хипотези за установяване на статистическа значимост.*

Използваната при статистическите заключения информация от наблюдение на извадки има известни ограничения по отношение на надеждността, точността и валидността, но въпреки това тя представлява основата за изграждане на медицинското познание, което имаме за човешките популации.

Всички статистически заключения изхождат от предпоставката, че ***една добра извадка трябва да бъде:***

* подбрана на основата на случаен непреднамерен подбор, за да се намали вероятността за систематична грешка;
* *представителна, за да се подобри валидността ѝ;*
* *достатъчно голяма по обем, за да се повиши точността на изчисляваните описателни характеристики.*

2. Същност на статистическото оценяване

Оценката се заключава в ***използване на резултатите от проучване върху сравнително малка извадка като мярка или индикация за нивата на съответните показатели в много по-широка популация***.

Например, количествените променливи в извадката най-често се описват чрез средна величина $\overline{x}$ и стандартно отклонение **s**. По подобен начин могат да бъдат описани и популациите като се използват символите: средна величина **μ** и стандартно отклонение σ.

Именно поради това, че стойностите на параметрите в популацията са неизвестни, най-напред се подбират извадки, установяват се стойностите на извадковите статистики и на тази основа се извличат статистически заключения за параметрите в популацията. Например, при измерване на диастолното налягане при 56 мъже пушачи на възраст 40–59 г. са получени $\overline{x}$ = 86 мм Hg и ***s*** = 14 мм

Hg. Бихме искали да знаем какво е средното ниво на диастолното кръвно налягане при всички мъже пушачи на тази възраст.

Оценката на даден параметър в популацията може да се извърши по два начина: ***точково и интервално оценяване.***

Точковото оценяване предоставя оценка на даден популационен параметър чрез една единствена стойност, която той най-вероятно може да приеме и най-често се изразява чрез извадковите статистики (напр. средна аритметична, пропорция и др.). Точковата оценка игнорира допусканата при извадковите проучвания репрезентативна грешка. Поради това тя не може да се разглежда като точна стойност на популационния параметър и няма самостоятелно приложение, а служи за основа на интервалната оценка.

Интервалното оценяване се опира на множество стойности, съсредоточени около точковата оценка, които формират определен интервал, в границите на който при определено ниво на гаранционна вероятност се предполага, че се намира истинската стойност на параметъра в популацията.

3. Основни понятия при статистическото оценяване

Преди да бъде направено заключение за параметрите на популацията на базата на статистиките от извадката трябва да изясним същността на следните основни понятия:

* *стандартна (средна стохастична) грешка;*
* *гаранционна вероятност (доверителност);*
* *гаранционен (доверителен) коефициент;*
* *максимална стохастична грешка*
* *интервал на доверителност.*

3.1. Стандартна (средна стохастична) грешка

Стандартната грешка измерва стандартното отклонение на разпределението на дадена извадкова статистика. Най-често се изчислява стандартната грешка на средната аритметична величи-

на Ако от една и съща популация са извлечени няколко извадки и е изчислена средната аритметична величина за всяка извадка, то стандартното отклонение на получените средни величини се нарича ***стандартна грешка на средната величина.*** Като такава тя характеризира точността на изчислените описателни статистики в извадката при използването им като оценъчни индикатори за параметрите в популацията. На практика, стандартната грешка се изчислява по данни само от една репрезентативна извадка и размерът на грешката може да се контролира чрез повлияване на определящите я фактори.

От какво зависи величината на стандартната грешка?

На първо място, вариабилността на дадена статистика в извадката (напр. за средната величина) зависи ***от варирането на индивидуалните наблюдения*****,** от които е композирана извадката. Колкото повече критерии и изисквания за еднородност на извадката са заложени при нейното сформиране, толкова по-малка е вариабилността, т. е. средната величина е съпроводена с по-малко стандартно отклонение.

На второ място, вариабилността на средната величина в извадката зависи ***от размера на извадката.*** Например, средният ръст на мъже студенти– медици, изчислен върху 100 наблюдения ще бъде много по-точен от този, изчислен върху 15 наблюдавани случая.

Следователно, ***вариабилността на средната величина в извадката нараства с индивидуалното вариране и намалява с увеличаване на размера на извадката.*** Това се вгражда във формулата за стандартната грешка по следния начин:

(10.1) $$s_{\bar{x}} = \frac{s}{\sqrt{n}},$$

където

$s_{\bar{x}}$ – стандартна грешка на средната величина в извадката

s – стандартно отклонение на средната величина в извадката

n – брой наблюдавани случай в извадката

Следователно, стандартната грешка на извадковата статистика е право пропорционална на стандартното отклонение и обратно пропорционална на размера на извадката. Стандартната грешка винаги ще бъде по-малка от самото стандартно отклонение, тъй като тя се получава като отношение на стандартното отклонение и квадратен корен от размера на извадката, а размерът на извадката винаги е по-голям от единица, т.е. квадратният корен винаги ще е по-голям от единица.

3.2. Гаранционна вероятност (доверителност)

Под ***гаранционна вероятност (доверителност)*** в широк смисъл на понятието се разбира ***вероятността, с която се подкрепя дадено твърдение.***

Статистическото заключение за параметрите в популацията винаги има вероятностен характер. Следователно, *в контекста на интервалното оценяване,* ***гаранционната вероятност е вероятността, с която се подкрепя твърдението, че оценяваният параметър от популацията попада в съответни доверителни граници.***

При оценката на резултатите от репрезентативните проучвания не е нужно да знаем как се изчисляват вероятностите, но е полезно да знаем връзката между честотните разпределения и вероятността.

Има редица стандартни разпределения, от които най-често срещано е нормалното разпределение. При него 68,2% от случаите прилягат най-близко до средната стойност – в интервала $\bar{x} \pm$ ***s,*** т.е. има 68,2% вероятност за тези стойности да се разполагат около средното ниво. По същия начин разсъждаваме за случаите, намиращи се на две стандартни отклонения вляво или вдясно от средната стойност, за които вероятността е 95% и т. н.

3.3. Гаранционен (доверителен) коефициент

Нивото на вероятност (в %) при нормалното разпределение съответства на точно определена числена стойност на ***t-критерия на Стюдент***, поради което той се нарича ***гаранционен (доверителен) коефициент***. При голям брой случаи (над 120) съотношението е следното (**табл. 10.1**):

Табл. 10.1. Връзка между стойността на t-критерия и гаранционната вероятност

Стойност на *t*-критерий	Вероятност в %
0.5	38.2%
1.00	68.2%
1.64	90.0%
1.96	95.0%
2.58	99.0%
3.29	99.999%

В медицината и здравеопазването статистическите заключения следва да бъдат подкрепяни с ***високо ниво на гаранционна вероятност*** – не по-малка от 95%, а в някои случаи, когато става дума за гранични области между живота и смъртта – дори над 99%. Стойността на доверителния коефициент t в тези случаи при извадка над 120 случая ще бъде приблизително равна на 2, а при 99% – над 2.5.

При размер на извадката под 120 случая стойността на доверителния коефициент се определя по специална таблица за критичните стойности на t-критерия (***Приложение 1***). При едно и също ниво на гаранционна вероятност стойността на t намалява с увеличаване на размера на извадката, а при един и същ размер на извадката t нараства с увеличаване на нивото на гаранционната вероятност.

3.4. Максимална стохастична грешка

Максималната стохастична грешка характеризира максималното отклонение на стандартната грешка от истинската стойност на параметъра и се определя като произведение на стандартната грешка и доверителния коефициент. Най-често се означава със символа Δ и за средната аритметична величина $\Delta = t.s_{\bar{x}}$

Практически максималната стохастична грешка се използва при изчисляване интервала на доверителност, тъй като коефициентът t е свързван пряко с възприетото ниво на гаранционна вероятност.

3.5. Интервал на доверителност (доверителни граници)

Както подчертахме, при статистическото оценяване на параметрите в популацията се използва вероятностен подход, който се свежда до определяне на ***интервал на доверителност (доверителни граници).***

Доверителният интервал (означава се с **ДИ или CI** – от Confidence Interval) ***представлява интервал, в границите на който при възприетата от изследователя гаранционна вероятност се намира истинската стойност на параметъра за популацията.***

Доверителният интервал се построява като към точковата оценка на извадковата статистика се прибави и извади максималната стохастична грешка. Последната включва в себе си гаранционната вероятност чрез стойността на доверителния коефициент ***t*** и по такъв начин се оформят долната и горната граница на интервала. Например, за средната аритметична величина в популацията доверителният интервал ще има следния вид:

(10.2) $\mathbf{CI} = \mu_1 \div \mu_2 = \bar{x} \pm \Delta = \bar{x} \pm t.\, s_{\bar{x}}$ **,**

където: $\mu_1 = \bar{x} - \Delta$ и $\mu_2 = \bar{x} + \Delta$

Следователно, истинската стойност на съответния параметър за популацията ще се намира в пределите на стойността на оценъчния индикатор за извадката плюс минус максималната стохастична грешка, т.е. μ ще бъде не по-малка от μ_1 и не по-голяма от μ_2**.**

Гаранционната вероятност на статистическия извод не зависи от размера на извадката – тя се задава предварително от изследователя, но при едно и също ниво на вероятност интервалната оценка ще бъде по-точна (интервалът ще бъде по-тесен), ако размерът на извадката е по-голям.

При един и същ масив от данни, ширината на доверителния интервал зависи от стойността на *t*-критерия. При $P = 95\%$ стойността на *t*-критерия е по-малка, отколкото при 99%, а оттук и по-тесен доверителен интервал. Обратно, при по-висока гаранционна вероятност интервалът на доверителност ще бъде по-широк, тъй като стойността на *t* нараства (при равни други условия – напр. един и същ размер на извадката и стандартното отклонение).

3.6. Степен на свобода

Терминът ***„степен на свобода“*** (означава се с ***„к“*** или ***df*** – от degree of freedom) за дадена променлива величина се използва за характеристика на броя на независимите части от информацията, съдържаща се в дадена статистика, т.е. *df* представлява брой резултати, които могат свободно да варират при изчисляването на дадена статистика, така че да не се промени крайния резултат. Например, ако в случайна извадка от n наблюдения сме изчислили средната аритметична, то за да получим същия резултат $n - 1$ от измерванията могат да варират, но последното измерване трябва да има такава стойност, че да не се промени сумата от направените измервания.

Определянето на степените на свобода е задължително условие при работа с таблиците за критичните стойности на *t*-критерия при интервално оценяване, както и при проверка на хипотези чрез параметрични и непараметрични методи и при други статистически методи на анализ.

4. Практически стъпки при статистическото оценяване

4.1. Оценка на средни величини

Обобщаването на резултатите за популацията на основата на наблюдения на извадка преминава през следните етапи:

I етап – Определяне на стандартната грешка

В примера в глава 5 за диастолното налягане в извадка от 56 мъже силни пушачи на възраст 40–59 г. е $\overline{x} = 86$ мм Hg и s = 14 мм. Стандартната грешка се определя по формулата:

(10.3) $$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{14}{\sqrt{56}} = \frac{14}{7{,}42} = \pm 1{,}8 \text{ mm Hg}$$

II етап – Избор на стойност на t-критерия съобразно броя на случаите в извадката и нивото на гаранционна вероятност, с която желаем да бъдат подкрепени нашите изводи.

При 56 наблюдения (степен на свобода $df = 56 - 1 = 55$) и гаранционна вероятност $P = 95\%$ от таблицата за критичните стойности на *t*-критерия (***Приложение 1***) намираме $t = 2.00$ (в реда съответстващ на $df = 60$).

III етап – Определяне на CI за средната величина в популацията

(10.4) $$\mathbf{CI} = \mu_1 \div \mu_2 = \overline{x} \pm \Delta = \overline{x} \pm \mathbf{t}.\, s_{\bar{x}},$$

където $\mu_1 = \overline{x} - \Delta$ и $\mu_2 = \overline{x} + \Delta$

Изчисляваме $\mu_1 = 86 - 2.00 \times 1.8 = 82.4$ mm Hg и
$\mu_2 = 86 + 2.00 \times 1.8 = 89.6$ mm Hg

Построяваме доверителния интервал
$\mathbf{CI} = \mu_1 \div \mu_2 = 82.4 \div 89.6$.

IV етап – Формулиране на заключение за стойността на параметъра в популацията

На основата на посочения пример може да се направи извод, че средната стойност на диастолното налягане за цялата популация (мъже, силни пушачи, на възраст 40–59 г.) се очаква да бъде не по-ниска от 82.4 и не по-висока от 89.6 мм Hg; това твърдение е подкрепено с гаранционна вероятност 95%.

Увеличаването на гаранционната вероятност на изводите за популацията изисква приемане на по-висока стойност на *t*-критерия, а следователно и ширината на интервала на доверителност ще се увеличи. Последното е напълно логично, тъй като ще намалее вероятността за неточност.

Когато средното ниво в извадката е представено чрез ***медиана*** ***(M_0),*** доверителният интервал се определя с помощта на специална таблица според броя на наблюдаваните случаи.

4.2. Оценка на коефициенти и пропорции

Оценката на коефициенти за честота и пропорции се различава само по формулата за стандартната грешка. Използват се следните символи:

	Извадка ***Известна*** ***Статистики***	***Популация*** ***Неизвестна*** ***Параметри***
Коефициенти/пропорции	***p***	π
Стандартно отклонение	***s***	σ

Преминава се през същите етапи на работа:

I етап – Определяне на стандартната грешка:

(10.5) $$s_p = \sqrt{p.q/n}\ ,$$

където:

s_p е стандартната грешка на коефициента (интензивния показател) или на пропорцията (екстензивния показател);
p – изчисленият от извадката показател;
q – противоположното до 1 или 100%, $q = 1 - p$ или $100 - p$

II етап – Определяне на величината на t-критерия според желаното от изследователя ниво на гаранционна вероятност на изводите и степента на свобода за извадката (по таблицата за критичните стойности на t-критерия)

III етап – Определяне на интервала на доверителност за параметъра в популацията –

(10.6) $$CI = \pi_1 \div \pi_2 = p \pm t.sp\ ,$$

където $\pi_1 = p - t.sp$ и $\pi_2 = p + t.sp$

IV етап – Формулиране на заключението за стойността на параметъра в популацията

5. Определяне на минималния размер на извадката за оценка на параметрите в популацията

Минималният размер на извадката зависи от:

* ***целта и постановката на проучването;***
* ***плана за статистически анализ*** (методи за статистическа обработка);
* ***точността на измерванията***, които трябва да бъдат направени;
* ***степента на точност при обобщаване на данните за популацията,*** т.е. допусканата стандартна грешка;
* ***гаранционната вероятност (доверителност)*** на заключенията

В практически план, изследователите задават предварително желаното ниво на гаранционната вероятност на статистическото заключение за популацията. Определя се предварително чрез пилотно проучване или по литературни данни очакваното ниво на извадковата статистика (средна величина или пропорция) и стойността на допустимата стандартна грешка. Накрая, чрез преобразуване на формулата за максималната стохастична грешка, в която остава неизвестен само броят на случаите (n), се определя необходимият размер на извадката.

5.1. Определяне на размера на извадка при количествени променливи

При подбор на извадка за изучаване на количествена характеристика, изследователят трябва да посочи:

* допустимата максимална грешка (в абсолютни или относителни единици);
* стандартното отклонение на променливата величина в популацията;
* възприетата гаранционна вероятност на заключенията – обикновено 95%.

Пример: Изследовател иска да оцени средното ниво на хемоглобина в дадена общност. Предварителната информация, с която разполага е, че средната е около 150mg/l със стандартно отклонение 32mg/l. Ако допустимата максимална грешка е до 5mg/l, колко лица трябва да бъдат включени в проучването?

От формулите $\Delta = t \, . \, s_{\bar{x}}$ и $s_{\bar{x}} = s/\sqrt{n}$ намираме, че $\Delta = t. \, s/\sqrt{n}$

Оттук $n = t^2 \, . \, s^2/\Delta^2 = [(1{,}96)^2 \, . \, (32)^2]/(5)^2 = 157.4$ лица

Ако размерът на популацията, от която ще се подбере извадката, е известен (напр. 3000 души), то необходимия размер на извадката ще бъде:

$n = (1{,}96)^2 \, . \, (32)^2 / [(5)^2 + (1{,}96)^2 \, . \, (32)^2 / 3000] = 149.5$ лица; т.е. извадката трябва да включва поне 150 лица.

5.2. Определяне на размера на извадка при качествени променливи

При подбор на извадка за изучаване на качествена характеристика, изследователят трябва да посочи:

* приблизителната стойност на пропорцията (p);
* допустимата максимална грешка (в абсолютни или относителни единици);
* стандартното отклонение на променливата величина в популацията;
* възприетата гаранционна вероятност на изводите.

От формулите $\Delta = t. \, s_{\bar{x}}$ и $s_{\bar{x}} = \sqrt{p.q/n}$ намираме, че

$$\Delta = t \, . \sqrt{p.q/n}$$

Оттук $n = t^2 \, . \, p \, . \, q/\Delta^2$

Пример: $p = 26\%$ (0,26); $\Delta = 3\%$ (0,03); $P = 95\%$; $t = 1.96$

Тогава: $n = t^2 \, . \, p \, . \, q/\Delta^2 = (1.96)^2 \, . \, 0{,}26 \, . \, 0{,}74/(0{,}03)^2 = 821{,}2$ лица.

Ако размерът на популацията, от която ще се подбере извадката, е известен (напр. 3000 души), то необходимия размер на извадката ще се коригира по следния начин:

$n = 821{,}2 \, / \, (1 + 821{,}2/3000) = 644{,}7$ лица; т.е. извадката трябва да включва поне 645 лица. Ако стойността на p е неизвестна, приеме $p = 50\%$. Тогава и q ще бъде 50% и по такъв начин и стандартната грешка ще има максимална стойност. Определеният на тази основа обем на извадката няма опасност да е твърде малък или прекалено голям.