

VI. СТАТИСТИЧЕСКИ МЕТОДИ ЗА ОЦЕНКА НА ЗДРАВНИТЕ ТЕХНОЛОГИИ

Какво ще научим в тази глава

Както се убедихте дотук, ОЗТ трябва да събере, организира, анализира, интерпретира и накрая да предостави във вид, в който могат да се вземат решения, информация, която освен това е прецизна и детайлна. Това не е възможно без статистическата наука. Необходима ви е при екстраполацията на данни, когато имате информация само за 30-дневна преживяемост, а е необходимо да оцените период от 10 години. Статистиката е важна при определяне на достоверността на хипотези – всяка оценка на здравна технология е хипотеза: едната алтернатива е по-ефективна от другата. Как без статистиката ще разберете какви са причинните зависимости между явленията и процесите? Как например се влияе степента на определена странична реакция от дозата на медикамента? Как точно протичат тези изчисления, как сме сигурни, че полученият резултат не се дължи на случайност и още много ще научите в тази глава.

Въпроси, на които ще намерим отговор

1. Какво е приложението на статистиката в ОЗТ?
2. Кои са основните понятия в статистиката?
3. Как се проверяват хипотези?
4. Какво е p-value?
5. Какво е корелационен и регресионен анализ?

По своята същност статистическата наука включва множество методи за извличане на познания от опит – научни методи за събиране, обработка, представяне, анализ и интерпретация на данни, разкриване на зависимости и извличане на заключения от цифрови данни. Статистиката като научна дисциплина е фундаментално средство за изследователска работа както в областта на общественото здраве, така и в сферата на оценките на здравните технологии. Нуждата от статистически подход е отдавна призната в епидемиологията и общественото здраве, тъй като тези области се характеризират с множество масови явления в общности и популации.

ОСНОВНИ ПОНЯТИЯ И ПРИЛОЖЕНИЕ НА СТАТИСТИКАТА В ОЗТ

Приложението на статистическите методи в здравеопазването най-често е по повод на:

1. Критичен анализ и използване на първични и вторични данни, публикувани в научната литература. На практика всички публикувани материали в авторитетните медицински издания съдържат някаква форма на статистически анализ. Следователно за тяхното разбиране се изискват определени статистически познания, които позволяват да се оценят критично публикуваните данни, да се разграничат достоверните данни и да се направят изводи за практиката.

2. Изследване на варирането при биологичните обекти. Вариабилността е присъща и в отговор на биологични рискове. Тя винаги води до проява на различия, които могат да се дължат на действителен ефект, на случайно вариране или на двете. Изкуството на статистиката е да определи каква част от варирането може да се припише на случайността или на реалния ефект.

3. Определяне на структурата и дизайна на научните проучвания. Основните моменти, чрез които статистиката може да подпомогне планирането на научните изследвания, включват:

- Определяне на необходимия обем на извадката чрез изчисляване на вероятността за получаване на определен ефект при даден брой пациенти, с отчитане на вида на променливата величина и очакваната вариабилност.

- Разработване на подходящ подход за събиране на достоверни данни. Важно е избраните подходи да бъдат апробирани в пилотно проучване.

- Избор на репрезентативна извадка и подходяща контролна група.

- Избор на подходящ дизайн на проучването – наблюдателен или експериментален.

- Организиране на подходящи процедури за събиране на данни при клинични и лабораторни изследвания, епидемиологични проучвания и популационни изследвания с цел намаляване на вероятността от грешки.

– Описание на характеристиките на дадена група или ситуация чрез представяне на данните в таблици и графики и изчисляване на обобщаващи измерители.

– Анализирани на данните и извлечени на заключения чрез използване на понятията за вероятност.

Основното приложение на статистиката в сферата на оценяване на здравните технологии е свързано с характеристиките на популацията, рисковия профил, влиянието на фактори от околната среда върху здравето, изследване на динамиката в популацията и данните за разпределение на ресурсите в здравеопазването. Конкретният анализ на приложимите статистически подходи е обект на тази глава.

Основни понятия в статистиката

1. Стандартна (средна стохастична) грешка
2. Гаранционна вероятност (доверителност)
3. Гаранционен (доверителен) коефициент
4. Максимална стохастична грешка
5. Интервал на доверителност.

Практически стъпки в статистическото оценяване

- а) *Оценка на средни величини*
- б) *Оценка на коефициенти и пропорции*
- в) *Определяне на минималния размер на извадката за оценка на параметрите в популацията.*

1. Статистическо оценяване при екстраполация на данни

На практика всички клинични изследвания и последващи оценки на здравни технологии се реализират на база данни от определена извадка от популацията или таргетираните пациенти.

Процесът на извличане на заключения за характеристиките на популацията от фактите, получени при наблюдение на извадка, се нарича статистическо заключение, или генерализация на данните.

Статистическите заключения решават две основни задачи:

– Оценка на резултатите от наблюдение на извадка и извличане на изводи за параметрите на съответната популация чрез екстраполация на данните.

– Сравняване на резултати от наблюдение на две или повече извадки – представлява тестване 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 представлява основата за изграждане на оценките на здравните технологии и резултатите от тях на популационно ниво.

В голяма степен обективността на статистическото заключение зависи от качеството на извадката. Най-общо всяка извадка трябва да бъде:

- Подбрана въз основа на случаен непреднамерен подбор, за да се намали вероятността от системна грешка;
- Представителна, за да се подобри валидността;
- Достатъчно голяма по обем, за да се повиши точността на изчисляваните описателни характеристики.

Същността на статистическата оценка се заключава в използване на резултатите от изследване върху сравнително малка извадка като мярка или индикация за нивата на съответните показатели в много по-широка популация.

Основни параметри в статистиката

Обикновено качествените променливи в една извадка, които са установени при проведено проучване, се описват чрез средна величина (\bar{x}) и стандартно отклонение (s). Това са оценъчните индикатори на извадката, въз основа на които се изчисляват съответните неизвестни параметри на популационно ниво – средна величина (μ) и стандартно отклонение (σ).

За да бъде направено заключение за параметрите на популацията на базата на оценъчните индикатори на извадката, е необходимо да се изясни същността на основните понятия в статистическото оценяване при екстраполацията на данните – стандартна (средна стохастична) грешка, гаранционна вероятност (доверителност), гаранционен (доверителен) ко-

ефициент, максимална стохастична грешка и интервал на доверителност.

1. Стандартна (средна стохастична) грешка

Стандартната грешка измерва стандартното отклонение на разпределението на оценъчните индикатори (\bar{x} , s) на извадката. Тя се среща в около 50% от научните публикации. Използва се за данни, които са “нормално разпределени”, за да получим информация за това какво количество от данните са около средните стойности.

Най-често се изчислява стандартна грешка на средната аритметична величина на индикаторите, извлечени от няколко извадки (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.

Стандартната грешка на средната величина зависи от вариационността на средната величина. От своя страна вариационността на средната величина в извадката нараства с индивидуалното вариране и намалява с увеличаване размера на извадката.

$$S\bar{x} = \frac{S}{\sqrt{n}}, \text{ където}$$

$S\bar{x}$ – стандартна грешка на средната величина

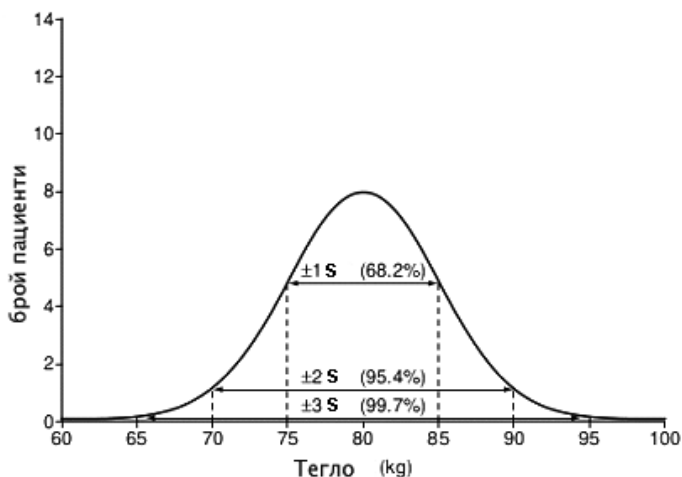
S – стандартно отклонение на средната величина в извадката

n – брой наблюдавани случаи в извадката.

Следователно стандартната грешка на средната величина е правопрпорционална на стандартното отклонение и обратно-пропорционална на размера на извадката.

Нека вземем за пример група пациенти, които са включени в проучване и имат нормално разпределение на теглото. Средното тегло (\bar{x}) на пациентите е 80 кг. За тази група изчисленото стандартно отклонение (s) е 5 кг.

Фиг 28. Графично представяне на средна величина (\bar{X}) и стандартно отклонение (s)



$\bar{X} \pm 1s$ – включва 68,2% от пациентите, които тежат между 75 и 85 кг

$\bar{X} \pm 2s$ – включва 95,4% от пациентите, които тежат между 70 и 90 кг

$\bar{X} \pm 3s$ – включва 99,7% от пациентите, които тежат между 65 и 95 кг.

Как се изчислява стандартна грешка?

Изчислението ще бъде илюстрирано с един пример

Ако измерим височината на 5-има ученици и получим следните резултати:

Име	Височина, измерена с точност до 0.5 см	Индивидуално отклонение ($\bar{x} - i$)	Отклонение на квадрат ($\bar{x} - i$) ²
1. Влади	150.5	11.9	141.61
2. Филип	170.0	-7.6	57.76
3. Христо	160.0	2.4	5.76
4. Краси	161.0	1.4	1.96
5. Росен	170.5	-8.1	65.61
n= 5 (размер на извадката)	средна величина $\bar{x} = 162.4$ cm		Сума от отклоненията на квадрат $\sum (\bar{x} - i)^2 = 272.70$

1. Можем да изчислим средната величина, като разделим сбора от всички измервания на броя на измерванията.

2. Можем да изчислим индивидуалното отклонение на всеки, като извадим неговата стойност от средната величина.

3. Повдига се индивидуалното отклонение на квадрат и така всички стойности ще станат положителни.

4. Сумират се всички квадрати на индивидуални отклонения.

5. Получената сума се разделя на размера на извадката минус едно. $\sum (\bar{x} - i)^2 / (n-1) = 272.70 / 4 = 68.175$

6. Изчислява се корен квадратен на полученото при предишната стъпка число: $\sum \sqrt{(\bar{x} - i)^2 / (n-1)} = \sqrt{68.175} = 8.257$ Така се получава стандартното отклонение (S).

7. Разделете стандартното отклонение на квадратния корен на размера на извадката, за да получите стандартната грешка $(S \bar{x}) = S / \sqrt{n} = 8.257 / 2.236 = 3.69$

2. Гаранционна вероятност (доверителност)

Под гаранционна вероятност се разбира вероятността, с която се подкрепя дадено твърдение. Статистическото заключение за параметрите в популацията винаги има вероятностен характер.

Гаранционната вероятност най-често се изразява чрез процент между 0 и 100% със значения на крайните точки – невъзможно и сигурно. Означава се най-често с „P“.

При оценка на резултатите от клиничните изпитвания или икономически оценки на здравните технологии е необходимо да се знае връзката между честотните разпределения и вероятността.

3. Гаранционен (доверителен) коефициент

Нивото на вероятност в % при нормалното разпределение съответства на точно определена числена стойност на t-критерия на Student, поради което той се нарича гаранционен (доверителен) коефициент. В табл. 14 е представено съотношението на вероятността спрямо t-критерия при извадки, съдържащи над 120 случая.

Таблица 14. t-критерий на Student при извадки над 120 случая

Стойност на t-критерия	Вероятност (%)
0,50	38,2
1,00	68,2
1,64	90,0
1,96	95,0
2,58	99,0
3,29	99,999

Източник: Медицинска статистика, 2011

При оценяването на здравните технологии на популационно ниво статистическите заключения трябва да бъдат подкрепени с високо ниво на гаранционна вероятност, не по-малка от 95% ($t \sim 2.0$). При размер на извадката под 120 случая стойността на доверителния коефициент се определя по специална таблица за критичните стойности на t-критерия (вж. Hassard, T., 1991).

4. Максимална стохастична грешка

Представлява максималното отклонение на стандартната грешка от истинската стойност на параметъра и се определя като произведение на стандартната грешка и доверителния коефициент.

$$\Delta = t \cdot S \bar{x}, \text{ където}$$

Δ – максимална стохастична грешка

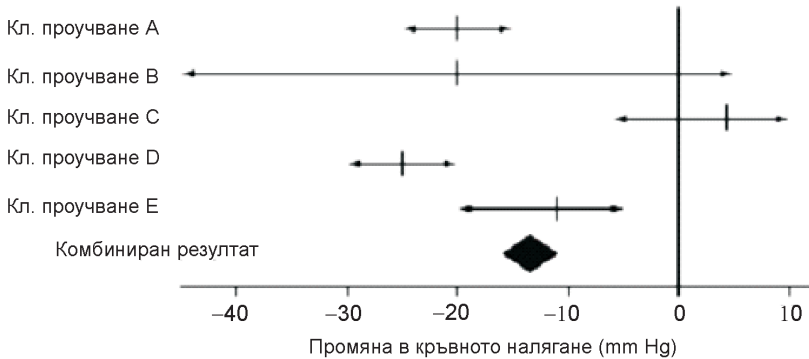
t – критерий на Student.

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

5. Интервал на доверителност

Доверителният интервал (CI) представлява интервал, в границите на който при възприетата от изследователя гаранционна вероятност се намира истинската стойност на параметъра за популацията при екстраполацията на данните от извадката. Графичен пример за изобразяване на интервал на доверителност е представен на фиг. 29.

Фигура. 29. Пример за интервал на доверителност



Използва се в около 75% от научните публикации. Най-просто казано, CI се използва когато вместо просто да изчислим средната стойност за дадена извадка, имаме нужда от по-голям диапазон, в който най-вероятно се съдържа стойността от реалната извадка, която да включва цялото население.

$CI = \mu_1 \div \mu_2 = (\bar{x} - \Delta) \div (\bar{x} + \Delta) = (\bar{x} - t \cdot S\bar{x}) \div (\bar{x} + t \cdot S\bar{x})$, където:

CI – доверителен интервал на средна величина за популацията

μ_1 – долна граница на средната аритметична величина

μ_2 – горна граница на средната аритметична величина.

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. При един и същ масив от данни, ширината на доверителния интервал зависи правопрпорционално от стойността на t-критерия.

При вероятност $P = 95\%$ стойността на t е по-малка, отколкото при вероятност $P = 99\%$, следователно при $P = 99\%$ ще имаме по-голям доверителен интервал.

Нека вземем за пример клинично проучване, което обхваща 100 пациенти, при които средното систолно кръвно налягане е 170 mm Hg. След приложение на изследван медикамент, средното систолно кръвно налягане спада с 20 mm Hg.

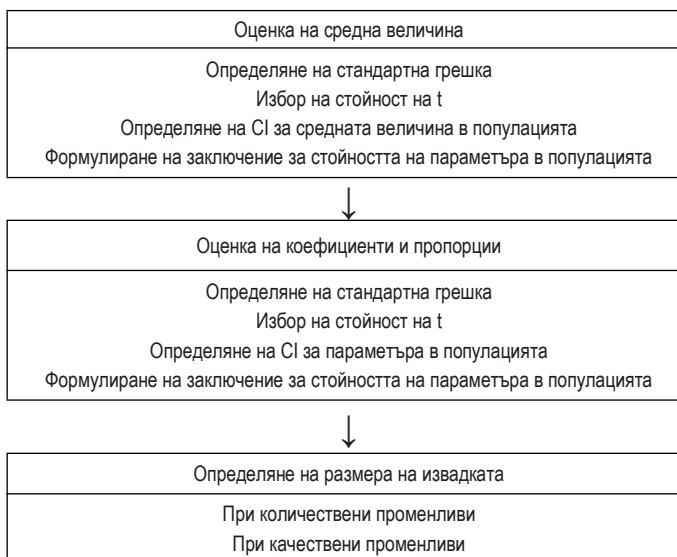
Ако 95%CI е 15-25 mm Hg, това означава, че можем да бъдем 95% убедени, че действителният ефект на медикамента е да понижи налягането с 15-25 mm Hg. Размерът на CI е в зависимост от размера на извадката. По-големите проучвания обикновено имат по-тесни интервали на доверителност – виж примера за графично изобразяване на интервалите на доверителност. На фигурата виждаме пет различни клинични проучвания на един и същи медикамент. Ординатната ос няма стойност и единствено служи да отбелязва къде е 0 (нула). Статистиците са обединили информацията от всичките пет проучвания и са изчислили, че общото средно понижение на кръвното налягане е 14 mm Hg, CI 12-16. Това е пример как като се комбинират данните от няколко проучвания, се получава много по-точен резултат.

КАК ДА ИЗВЪРШИМ ЕДНА СТАТИСТИЧЕСКА ОЦЕНКА?

Практически етапи в статистическото оценяване

След разглеждане на основните понятия и параметри при статистическото оценяване трябва да се анализират практическите етапи на реализирането му (фиг. 30).

Фигура 30. Етапен алгоритъм за статистическо оценяване



а) Оценка на средни величини

Както вече отбелязахме, средната величина \bar{x} представлява средна аритметична величина на разглежданите данни:

$$\bar{x} = \frac{\sum_{i=1}^n x}{n}$$

– Определяне на стандартната грешка

Разглеждаме хипотетичен пример за влиянието на тютюнопушенето върху диастолното кръвно налягане при 224 мъже на възраст между 50 и 60 год. – табл. 15.

Таблица 15. Стойност на диастолното налягане при мъже пушачи между 50 и 60 год.

Диастолно налягане, mm Hg	Абсолютна честота, бр.	Относителна честота, %	Кумулативна честота, %
60-64,9	4	1,8	–
65-69,9	4	1,8	3,6
70-74,9	8	3,6	7,2
75-79,9	20	8,9	16,1
80-84,9	36	16,1	32,2
85-89,9	40	17,8	50,0
90-94,9	44	19,6	69,6
95-99,9	28	12,5	82,1
100-104,9	20	8,9	91,0
105-109,9	12	5,4	96,4
110-114,9	4	1,8	98,2
115-119,9	4	1,8	100,00
	N = 224	100,00	

В разглеждания хипотетичен пример средната величина е $\bar{x} = 86$ mm Hg и стандартното отклонение е $S = 14$ mm Hg.

Следователно стандартната грешка на средната величина ще бъде:

$$S\bar{x} = \frac{S}{\sqrt{n}} = \frac{14}{\sqrt{224}} = \frac{14}{14.97} = \pm 0.94 \text{ mm Hg}$$

- Избор на стойност на t-критерия

При $n = 224$ и гаранционна вероятност $P = 95\%$ от табл. 15 за t-критерия на Student, намираме $t = 1.96$.

- Определяне на CI за средната величина в популацията

$$CI = \mu_1 \div \mu_2$$

$$\mu_1 = \bar{x} - \Delta = \bar{x} - t \cdot S\bar{x} = 86 - 1,96 \cdot 0,94 = 84,16 \text{ mm Hg}$$

$$\mu_2 = \bar{x} + \Delta = \bar{x} + t \cdot S\bar{x} = 86 + 1,96 \cdot 0,94 = 87,84 \text{ mm Hg}$$

Следователно доверителният интервал е $CI = 84,16 \div 87,84$

- Формулиране на заключение за стойността на параметъра в популацията. На основание на посочения пример може да се направи статистически извод, че средната стойност на диастолното налягане за цялата популация в риск (мъже пушачи на възраст между 50 и 60 год.) се очаква да бъде не по-ниска от 84,16 mm Hg и не по-висока от 89,6 mm Hg. Това твърдение е подкрепено с гаранционна вероятност 95%. Или казано по друг начин, можем с 95% сигурност да твърдим, че в цялата популация диастолното налягане е между 84,16 mm Hg и 89,6 mm Hg.

б) Оценка на коефициенти и пропорции

Оценката на коефициенти за честота и пропорции се различава само по формулата за стандартна грешка.

- Определяне на стандартната грешка

$$Sp = \sqrt{p \cdot q/n}, \text{ където:}$$

Sp – стандартна грешка на коефициента (интензивния показател) или на пропорцията (екстензивния показател) P – изчисленият показател от извадката

$$q = 1 - p \text{ или } 100\% - p$$

- Определяне на величината на t-критерия (по табл. за t-критерия на Student)

- Определяне на интервала на доверителност за параметъра в популацията

$$CI = \pi_1 \pm \pi_2 = p \pm t \cdot Sp, \text{ където } \pi_1 = p - t \cdot Sp \text{ и } \pi_2 = p + t \cdot Sp$$

- Формулиране на заключението за стойността на параметъра в популацията

в) 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 (методите за статистическа обработка), точността на измерванията, които трябва да бъдат направени, степента на точност при обобщаване на данните за популацията, допусканата стандартна грешка и гаранционната вероятност (доверителност) на заключенията.

В практически план изследователите задават предварително желаното ниво на гаранционна вероятност на статистическото заключение за популацията (напр. $P = 95\%$). 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.

- Определяне на размера на извадка при количествени променливи

Разглеждаме хипотетична задача за оценяване на средното ниво на диастолното кръвно налягане в определена таргетна популация. Предварителната информация, с която се разполага по литературни данни, е, че средното ниво е около 86 mm Hg със стандартно отклонение 14 mm Hg. Ако допустимата максимална грешка е 2 mm Hg, колко лица трябва да бъдат включени в проучването?

$$\Delta = t \cdot S\bar{x}; S\bar{x} = S/\sqrt{n} \Rightarrow \Delta = t \cdot S/\sqrt{n} \Rightarrow n = t^2 \cdot S^2/\Delta^2 = (1,96^2 \cdot 14^2)/2^2 = 188,24 \text{ лица.}$$

Ако размерът на популацията, от която ще се избере извадката, е известен (напр. $p = 5000$ души), то необходимият размер на извадката ще бъде:

$$n = t^2 \cdot S^2 / (\Delta^2 + t^2 \cdot S^2/p) = (1,96^2 \cdot 14^2) / (2^2 + 1,96^2 \cdot 14^2/5000) = 181,41 \text{ лица}$$

– Определяне на размера на извадката при качествени променливи

При подбор на извадка за изучаване на качествена характеристика изследователят трябва да посочи приблизителната стойност на пропорцията (p), допустимата максимална грешка (Δ), стандартното отклонение на променливата в популацията (s) и възприетата гаранционна вероятност на изводите (напр. $P = 95\%$).

$$\Delta = t \cdot Sp; Sp = \sqrt{p} \cdot q/n \Rightarrow \Delta = t \cdot \sqrt{p} \cdot q/n \Rightarrow n = t^2 \cdot p \cdot q / \Delta^2$$

Пример: $p = 38\%$ (0,38), $\Delta = 3\%$ (0,03), $t = 1,96$, $q = 1 - p = 1 - 0,38 = 0,62$

$$n = t^2 \cdot p \cdot q / \Delta^2 = 1,96^2 \cdot 0,38 \cdot 0,62 / 0,03^2 = 1005,65 \text{ лица}$$

2. Статистически подходи за проверка на хипотези

Много често изследователските стратегии включват събиране на данни за целите на проверката на определени хипотези. При оценката на здравните технологии почти винаги се провеждат изследвания върху ефективността на различни терапевтични подходи при лечение на пациенти с аналогични заболявания. Проучва се влиянието на определени фактори за възникването и развитието на съответни заболявания, като се използват опитни и контролни групи или се сравняват резултатите в едни и същи групи преди и след провеждане на съответната интервенция. Естествено възникват редица въпроси доколко наблюдаваните различия при сравнителните изследвания в извадките могат да се възприемат за достоверни, дължащи се на закономерни фактори, или те са резултат на случайност. Отговорът на тези въпроси се дава чрез сравняване на резултатите в наблюдаваните извадки с помощта на различни тестове за значимост, които позволяват на изследователя да установи дали данните подкрепят или опровергават формулираната от него изследователска хипотеза. Използваната статистическа процедура се нарича проверка (тестване) на хипотеза.

За да формулира хипотези за различията между групите и за взаимовръзките между променливите величини, изследователят трябва добре да познава публикуваните литературни данни по проблема, да извърши критичен анализ на данните (вж. гл. VIII), да избере подходящ дизайн на проучването и методи за събиране на достоверна информация.

Проверката на хипотези се основава на статистически тестове, които позволяват да се определи дали различията между оценъчните индикатори от извадката и параметрите на популацията или различията между статистиките на анализиранияте извадки се дължат на случайност, или са значими, закономерно обусловени от конкретни фактори и причини.

Формулирането и проверката на хипотези представляват съществена част от статистическите изводи и заключения и от процеса на оценяване на здравните технологии.

Когато се сравняват две здравни технологии или лекарствени терапии, обикновено се приема „нулева хипотеза“ (H_0), която най-общо гласи, че новата здравна технология не е по-ефективна от съществуващата терапевтична алтернатива, определена като „златен стандарт“. На H_0 се отдава специално внимание, защото тя се отнася до твърдението, което трябва да се провери. Твърдението, че новата здравна технология е по-ефективна от съществуващата алтернатива „златен стандарт“, се определя като „алтернативна хипотеза“ (H_1). Следователно H_1 се отнася до твърдението, което трябва да бъде прието, ако H_0 се отхвърли.

Окончателното заключение след прилагане на теста за проверка на хипотези винаги се дава от гледна точка на H_0 и може да бъде „ H_0 се отхвърля в полза на H_1 “ или „ H_0 не се отхвърля“.

Грешките, които могат да бъдат допуснати при проверката на хипотези, също се определят от гледна точка на нулевата хипотеза (H_0). Могат да бъдат направени две потенциални грешки – грешка от I род и грешка от II род.

Грешка от I род се получава, ако H_0 се отхвърли, когато в действителност е вярна. Вероятността за грешки от I род (α) може да бъде точно изчислена и намалена чрез промяна на ни-

вото на значимост на хипотезата. Обикновено се работи с ниво на вероятност не по-малко от 95%. Следователно H_0 се приема за вярна само ако нейната значимост е $> 0,05$. Ако приемем ниво на значимост 0,01, вместо 0,05, тогава възможността за допускане на грешка от I род намалява 5 пъти.

Грешка от II род се получава, когато H_0 не се отхвърли, когато в действителност е вярна H_1 (т.е. новата здравна технология е по-ефективна от съществуващата алтернатива). Вероятността за грешка от II род (β) обикновено е неизвестна и най-често се дължи на малки по размер извадки. Следователно грешка от II род може да се избегне чрез увеличаване размера на извадката.

Двата вида грешки са тясно и обратнопропорционално свързани – намаляването на α увеличава β и обратното.

Един от най-съществените въпроси при оценяване на здравните технологии се свежда до това кой вид грешка е по-допустима за изследователя и от каква гледна точка – на пациентите, на фондовете, които плащат, или на доставчика на новата здравна технология. Практиката сочи, че за повечето изследователи грешка от II род е недопустима, защото в масовия случай изследванията се спонсорират от доставчиците на новите здравни технологии, които се интересуват от хипотеза H_1 .

При анализа на данни от извадки изследователят най-често се стреми да установи има ли статистически значима връзка между променливите или статистически значима разлика между групите.

Понятието статистическа значимост е свързано с теорията на вероятностите. Нивото на значимост представлява вероятността, над която H_0 се приема за вярна и под която H_0 се отхвърля.

Това означава, че за да бъде един резултат статистически значим, не трябва да има повече от 5% вероятност той да се дължи на случайност.

Това е широко използваната т.нар. p-value. Не е необходимо да знаете как се изчислява, но трябва да можете да я интерпретирате. Колкото е по-ниска p-value, толкова по-малко вероятно е разликата да е в резултат на случайност. Това увеличава значимостта на установените резултати. $P = 0,01$ означава, че

вероятността за шанс е 1 на 100. Колкото е по-ниска стойността на p-value, толкова е по-малък шансът нулевата хипотеза да е достоверна.

Съществуват множество тестове, които се използват при проверка на хипотези по концептуално сходен начин, но изборът на подходящи статистически критерии за анализ на данните в конкретните изследвания се определя от следните основни съображения:

а) вида използвана скала за измерване на променливите величини в извадките (номинална, ординална, интервална, пропорционална);

б) броя на извадките (групите), които се включени в дадено изследване;

в) вида на сравняваните извадки – дали измерванията са получени от независими или от зависими извадки (повторни измервания при едни и същи пациенти).

Подходящият избор на статистически тестове за значимост е представен в табл. 16.

Таблица 16. Избор на статистически тестове за значимост при проверка на хипотези

Вид скала	Две извадки		Три и повече извадки	
	Независими	Зависими	Независими	Зависими
Номинална	Критерий на Pearson (χ^2)	Тест на McNemar	Критерий на Pearson (χ^2)	Тест на Cochran
Ординална	U-критерий на Mann-Whitney	Тест на Wilcoxon (знаков тест)	H-тест на Kruskal-Wallis (медианен тест)	Тест на Friedman
Интервална или пропорционална	t-критерий на Student	t-критерий на Student	F-критерий на Fisher (дисперсионен анализ)	F-критерий на Fisher (дисперсионен анализ)

Източник: Медицинска статистика, 2011

В зависимост от дизайна на конкретното изследване още преди събирането на данните трябва да се избере подходящ тест за анализа им.

МЕТОДИ ЗА ПРОВЕРКА НА ХИПОТЕЗИ

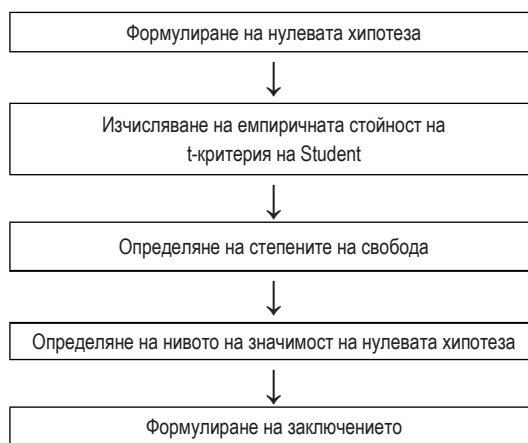
Методите за проверка на хипотези (статистически тестове) могат да бъдат класифицирани в две големи групи – параметрични и непараметрични.

1. Параметрични методи за проверка на хипотези

а) Сравняване на две наблюдавани групи (извадки)

– Сравняване на средни величини чрез t-критерий на Student. Това е един от най-често използваните тестове за значимост. Основава се на разликата между средните аритметични величини в извадките. Етапите в прилагането на този параметричен статистически тест са илюстрирани на фиг. 31.

Фигура 31. Етапен алгоритъм за сравняване на средни величини чрез t-критерия на Student



Разглеждаме хипотетичен пример за въвеждане на нов лекарствен продукт в кардиохирургично отделение, който цели да намали болничния престой след хирургично лечение на остър коронарен синдром. От наблюдение на 100 пациенти, изписани преди внедряването на новата здравна технология, е установена средна продължителност на престоя след операция $\bar{x}_1 = 30$ дни и стандартно отклонение $S_1 = 8$ дни. За група от 64 пациенти, при които е прилагана новата лекарствена терапия, средни-

ят престой е $\bar{x}_2 = 24$ дни и $S_2 = 6$ дни. Задачата на експерта по оценки на здравни технологии е да установи тези данни доказват ли, че новата лекарствена терапия оказва влияние върху следоперативния престой на болните в сравнение с използваната преди това терапия?

Първата стъпка е да се формулира нулевата хипотеза (H_0) – новата лекарствена терапия не оказва влияние върху следоперативния престой на пациентите. Допуска се, че разликите, които се наблюдават, са резултат на случайност. След това се изчислява емпиричната стойност на t-критерия на Student.

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{|30 - 24|}{\sqrt{\frac{8^2}{100} + \frac{6^2}{64}}} = \frac{6}{1,0966} = 5,47, \text{ където}$$

$|\bar{x}_1 - \bar{x}_2|$ – абсолютна стойност на разликата между средните величини в двете извадки

S_1, S_2 – стандартно отклонение (дисперсия) в двете извадки
 n_1, n_2 – брой пациенти в двете извадки.

Следващата стъпка е да се определят „степени на свобода“. Степените на свобода (dF) са свързани с броя на участниците и резултатите в определена група и с идеята за свобода на вариране. При проверка на хипотези с две извадки се използва формулата:

$$df = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2 = 100 + 64 - 2 = 162 = \infty$$

Следва определяне нивото на значимост на H_0 чрез сравняване на емпиричната стойност на t-критерия при съответната степен на свобода df в табл. 17 за критичните стойности на t-критерия ($n < 120$).

Таблица 17. Критични стойности на t-критерия на Student

dF	Ниво на значимост H_0 при двустранен тест						
	P = 0,1	P = 0,05	P = 0,02	P = 0,01	P = 0,005	P = 0,002	P = 0,001
	Ниво на значимост H_0 при едностранен тест						
	P = 0,05	P = 0,025	P = 0,01	P = 0,005	P = 0,0025	P = 0,001	P = 0,0005
1	6,314	12,706	31,821	63,657	127,32	318,31	636,62
2	2,92	4,303	6,965	9,925	14,089	22,327	31,598
3	2,353	3,182	4,541	5,841	7,453	10,214	12,924
4	2,132	2,776	3,747	4,604	5,498	7,173	8,610
5	2,015	2,571	3,365	4,032	4,773	5,893	6,869
6	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	1,753	2,131	2,602	2,947	3,268	3,733	4,073
16	1,764	2,120	2,583	2,921	3,252	3,686	4,015
17	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	1,714	2,069	2,500	2,807	3,104	3,485	3,767
24	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	1,703	2,052	2,473	2,771	3,057	3,421	3,690
28	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	1,699	2,045	2,464	2,756	3,038	3,396	3,659
30	1,697	2,042	2,457	2,750	3,030	3,385	3,646
40	1,684	2,021	2,423	2,704	2,971	3,307	3,551
60	1,671	2,000	2,390	2,660	2,915	3,232	3,460
120	1,658	1,980	2,358	2,617	2,860	3,160	3,373
∞	1,645	1,960	2,326	2,576	2,807	3,090	3,291

Източник: Hassard, T., Understanding Biostatistics, 1991

При $t = 5,47$ и $dF = \infty$ нивото на значимост H_0 при двустранен тест съответства на $P_{(t)} < 0,001$, защото $t = 5,47 > t = 3,291$, което съответства на $P_{(t)} = 0,001$.

Последният етап от сравнението е формулиране на заключението. Получената стойност $P_{(t)} < 0,001$ е значително отдалечена от предварително приетата критична стойност за правдоподобност на H_0 ($P = 0,05$ или $P = 0,01$). 1 на 1000 е вероятността резултатът да се дължи на шанс. Следователно H_0 не е вярна и се отхвърля.

Възприема се H_1 – в резултат на използването на новата лекарствена терапия има статистически значимо различие в средния следоперативен престой при сравняваните групи. Следователно новата здравна технология е по-ефективна от съществуващата алтернатива от гледна точка на следоперативния престой.

– Сравняване на коефициенти и пропорции чрез t-критерия на Student

Ако изчислените статистически стойности в наблюдаваните групи са пропорции или интензивни показатели, тогава t-критерият на Student се изчислява по следната формула:

$$t = \frac{|p_1 - p_2|}{\sqrt{\frac{p_1 \cdot q_1}{n_1} + \frac{p_2 \cdot q_2}{n_2}}}, \text{ където}$$

p_1, p_2 – изчислени статистически стойности в двете извадки (пропорция, коефициент за честота и др.)

q_1, q_2 – стойност, допълваща p_1, p_2 до 1, 10, 100, 1000 и т.н.

$q_1 = 100 - p_1$ или $q_1 = 1000 - p_1$

n_1, n_2 – брой наблюдавани пациенти в двете групи.

След определяне стойността на t-критерия и степените на свобода dF по таблицата за критичните стойности на t се намира $P_{(t)}$ и по същия начин се формулират изводите.

Ако $P > 0,05$, тогава H_0 е вярна и обратното.

– Определяне на необходим брой случаи (обем на извадките) при сравняване на две групи

При планиране на дадено проучване изследователят се нуждае от реалистична оценка за това колко лица са необходими за проверка на формулираната хипотеза. Изходният момент при сравняването и проверката на хипотези е колко лица са необходими, за да се установи значимо различие.

Разработени са множество таблици за определяне размера на извадката при различни методи: при работа с t-критерий, при дисперсионен и регресионен анализ и др. Когато в постановката на проучването се предвижда да се работи с две или повече групи, обемът на извадките може да се определи чрез преобразуване на формулите на сравняване на средни величини и пропорции.

б) Сравняване на повече от две групи чрез дисперсионен анализ (F-критерий на Fisher)

Доста често в изследователската практика се провеждат проучвания върху повече от две групи с цел да се определи дали съществуват достоверни различия между групите – например групи пациенти, лекувани с различни дози на един и същ лекарствен продукт. В този случай е препоръчително прилагането на дисперсионен статистически анализ, защото t-критерият на Student не може да се използва при повече от две групи.

Условията за прилагане на дисперсионен анализ са следните:

– Независимата променлива (фактор) се предпочита да бъде измерена върху номинална скала с две или повече нива. Например полът е номинална променлива с две нива (мъж, жена), докато социалният статус е номинална променлива с повече от две нива (работещ, безработен, учещ и т.н.)

– Зависимата променлива (резултат) да бъде непрекъснатата, измерена върху интервална или пропорционална скала и да има нормално разпределение.

– Сравняваните групи да бъдат взаимно изключващи се (независими една от друга) и да имат еднакви дисперсии (изискване за хомогенност на дисперсиите).

Съществуват няколко основни схеми за прилагане на дисперсионен анализ:

– В зависимост от броя на независимите променливи дисперсионният анализ може да бъде еднофакторен, двуфакторен или многофакторен.

– В зависимост от броя на участниците дисперсионният анализ може да бъде равномерен комплекс (еднакъв брой участници в групите) или неравномерен комплекс (различен брой участници в групите).

Статистическият въпрос при дисперсионния анализ се основава на допускане (нулева хипотеза), че всички изследвани групи са еднакви, извлечени от една и съща популация и всякакво наблюдавано различие в резултатите при сравняваните групи се дължи на случайност. Принципната разлика между проверката на хипотези с помощта на t-критерия на Student и дисперсионния анализ е в това, че при първия подход се сравняват директно средните аритметични величини в групите, а вторият подход се основава на сравняване на дисперсиите. Следователно същността на проверката на H_0 при дисперсионния анализ е свързана с обсъждане на варирането на резултатите.

Същността на дисперсионния анализ представлява измерване на следните параметри:

– Вътрегрупово вариране (дисперсия – S_i^2). Дължи се на индивидуалната вариабилност на наблюдаваните участъци в групата (пол, възраст, кръвно налягане, серумен холестерол и др.).

– Междугрупово вариране (дисперсия – S_e^2). Дължи се на различия в условията на провежданите експерименти и в нивата на изучаваните фактори. Следователно междугруповото вариране отразява влиянието на проучваните фактори или условията.

– Общо вариране (дисперсия – S_t^2). Представява сума от вътрегруповото и междугруповото вариране.

$$S_t^2 = S_i^2 + S_e^2$$

Отговорът на въпроса „Има ли значимо (сигнификантно) различие между резултатите в групите?“ се свежда до сравняване на междугруповото S_e^2 и вътрегруповото вариране (S_i^2).

Използва се формулата на F-критерий на Fisher.

$$F = S_0^2 / S_1^2$$

Ако не съществува реално различие между сравняваните групи, тогава стойността на F е близка до 1. Ако има сигнификантни различия между групите, тогава $F > 1$.

Изчисляването на дисперсиите S_0^2 и S_1^2 е сложен математически процес и в практиката се използват специализирани софтуерни продукти, като Statgraphics, SPSS и др.

2. Непараметрични методи за проверка на хипотези

Непараметричните методи са голяма група статистически критерии за проверка на хипотези. Те имат редица предимства пред параметричните методи:

- Параметричните методи се прилагат само при количествени променливи, докато непараметричните методи могат да се прилагат както за количествени, така и за качествени променливи.

- Непараметричните методи се използват при всички форми на разпределение на случаите.

- Непараметричните методи са по-икономични от гледна точка на ресурси.

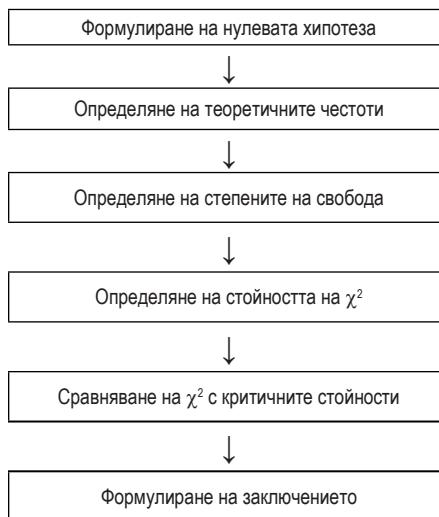
- Недостатък на непараметричните методи е по-малката значимост на използваните при тях статистически критерии.

а) Критерий на Pearson (χ^2)

Това е най-често използваният непараметричен тест за оценка на хипотези. Използва се в около 1/3 от научните публикации. Принципната основа на неговото приложение се състои в сравняване на фактическите честоти (f) с теоретичните честоти (f_t). Фактическите честоти (f) са тези, които са измерени в конкретното проучване, а теоретичните честоти (f_t) са тези, които биха се получили, ако е вярна нулевата хипотеза (H_0).

На фиг. 32 е представен етапният алгоритъм за изчисление на критерия на Pearson.

Фигура 32. Етапен алгоритъм за изчисление на критерия на Pearson (χ^2)



Логично следва изводът, че колкото е по-голяма разликата между f и f_t , толкова е по-голяма вероятността различието да бъде съществено и да се отхвърли H_0 , и обратното.

Изчисляването на χ^2 (кси) ще бъде илюстрирано с един хипотетичен пример, в който за 10-годишен период се наблюдават две групи мъже (пушачи и непушачи), като се проследява възникването на хронична обструктивна белодробна болест (ХОББ).

Данните са представени в четирикратна табл. 18, в която тютюнопушенето е независима променлива (проучван фактор), а ХОББ е зависима променлива (резултат).

Таблица 18. Изчисляване на критерий на Pearson (χ^2) при проследяване на зависимост между тютюнопушене и ХОББ

	Участници с ХОББ	Участници без ХОББ	Общо
Пушачи, $f(f_i)$	100 (75) a	400 (425) b	500 (a + b)
Непушачи, $f(f_i)$	50 (75) c	450 (425) d	500 (c + d)
Общо	150 (a + c)	850 (b + d)	1000

Целта на проучването е да се отговори на въпроса: Тютюн

о- пушенето оказва ли съществено влияние върху , f , f .
 заболяемостта от ХОББ? Първо се формулира нулевата хипотеза (H) – тютю- нопушенето не оказва влияние върху заболяемостта от ХОББ. Различията в заболяемостта при пушачи и непушачи са несъществени и се дължат на случайни фактори. Следващата стъп- ка е да се определят теоретичните честоти (f) във всяка клетка от таблицата.

$$f_t = \frac{\text{брой случаи в сумарен ред} \times \text{брой случаи в сумарна колона}}{\text{общ брой наблюдавани случаи}}$$

Следователно

$$f_{ta} = \frac{500 \cdot 150}{1000} = 75$$

Аналогично се изчисляват посочените в таблицата f_{tb}, f_{tc}, f_{td}. Следващата стъпка е определяне на степените на свобода (dF). Те се определят при този непараметричен метод по формулата: dF = (S-1) (R-1), където

S – брой на нивата на независимата променлива (фактор) R – брой на нивата на зависимата променлива (резултат). При четирикратна таблица, каквато е разглежданата в нашия хипотетичен пример, dF винаги има стойност 1, защото

$$dF = (2 - 1) (2 - 1) = 1$$

След това се изчислява стойността на χ^2 по формулата

$$\chi^2 = \sum \frac{(f - f_t)^2}{f_t}$$

¹ За четирикратна таблица формулата се опростява до:

26

$$\chi^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} = \frac{1000(100 \cdot 450 - 50 \cdot 400)^2}{500 \cdot 500 \cdot 150 \cdot 850} = 19,6$$

След това стойността на χ^2 се сравнява с критичните стойности на χ^2 – табл. 19.

Таблица 19. Критични стойности на критерия на Pearson (χ^2)

dF H_1	H_0	P = 0,100	P = 0,050	P = 0,025	P = 0,010	P = 0,005	P = 0,001
		1-P = 0,900	1-P = 0,950	1-P = 0,975	1-P = 0,990	1-P = 0,995	1-P = 0,999
1		2,71	3,84	5,02	6,63	7,88	10,83
2		4,61	5,99	7,38	9,21	10,60	13,82
3		6,25	7,81	9,35	11,34	12,84	16,27
4		7,78	9,49	11,14	13,28	14,86	18,47
5		9,24	11,07	12,83	15,09	16,75	20,52
6		10,64	12,59	14,45	16,81	18,55	22,46
7		12,02	14,07	16,01	18,48	20,28	24,32
8		13,36	15,51	17,53	20,49	21,96	26,13
9		14,68	16,92	19,02	21,67	23,59	27,88
10		15,99	18,31	20,48	23,21	25,19	29,59
11		17,28	19,68	21,92	24,73	26,76	31,26
12		18,55	21,03	23,34	26,22	28,30	32,91
13		19,81	22,36	24,74	27,69	29,82	34,53
14		21,06	23,68	26,12	29,14	31,32	36,12
15		22,31	25,00	27,49	30,58	32,80	37,70
16		23,54	26,30	28,85	32,00	34,27	39,25
17		24,77	27,59	30,19	33,41	35,72	40,79
18		25,99	28,87	31,53	34,81	37,16	42,31
19		27,20	30,14	32,85	36,19	38,58	43,88
20		28,41	31,41	34,17	37,57	40,00	45,32
21		29,62	32,67	35,48	38,93	41,40	46,80
22		30,81	33,92	36,78	40,29	42,80	48,27
23		30,81	33,92	36,78	40,29	42,80	48,27
24		33,20	36,42	39,36	42,98	45,56	51,18
25		34,38	37,65	40,65	44,31	46,93	52,62
26		35,56	38,89	41,92	45,64	48,29	54,05
27		36,74	40,11	43,19	46,96	49,64	55,48
28		37,92	41,34	44,46	48,28	50,99	56,89
29		39,09	42,56	45,72	49,59	52,34	58,30
30		40,26	43,77	46,98	50,89	53,67	59,70
40		51,81	55,76	59,34	63,89	66,77	73,40
50		93,17	67,50	71,42	76,15	79,49	86,66
60		74,40	79,08	83,30	88,38	91,95	99,61

Източник: Hassard, T., Understanding Biostatistics, 1991

По таблицата за критичните стойности на χ^2 в реда, съответстващ на dF = 1, намираме, че изчисленият $\chi^2 = 19,6$ се разполага вдясно от последната колона. Това означава, че възприетата H_0 има вероятност $P < 0,001$. Следователно H_0 е невярна и се отхвър-

ля. Възприема се хипотеза H_1 . Крайният извод се формулира по следния начин – има съществено различие в заболяемостта от ХОББ при пушачи и непушачи. Следователно тютюнопушенето оказва съществено влияние върху заболяемостта от ХОББ.

Тестът на Pearson е по-лесен за изчисление от статистиците, но предоставя само приблизителни данни за p-value и е неподходящ за малки извадки.

б) Други непараметрични критерии

– U-критерий на Mann-Whitney. Представлява един от най-мощните непараметрични тестове за сравняване на две популации на основата на две независими извадки. Използва се при 1 от 6 научни публикации.

U-критерият е аналог на t-критерия при сравняване на средни величини от независими извадки и се прилага, когато разпределението е неизвестно или се различава от нормалното. Изчисляването на емпиричната стойност на U-критерия се основава на прегрупиране на данните в извадките и подреждането им по рангове. За оценка на вероятността на H_0 също се използват специални таблици за U при еднакви или различни по размер извадки и с отчитане на посоката на теста – едностранен или двустранен.

– Медианен тест. Използва се при две независими извадки и се основава на сравняване на медианите. Данните се представят в рангова скала. За всяка извадка се определят медианата и броят случаи, които попадат под и над нея. След това се построява четирикратна таблица и се изчислява χ^2 , след което се сравнява с таблицата за критичните стойности на χ^2 . Медианният тест може да се прилага и при повече от две независими извадки с ординално измерване на зависимата променлива.

– Критерий на Kruskal-Wallis. Използва се при три и повече извадки. Представлява аналог на F-критерия на Fisher при еднофакторен дисперсионен анализ и е подходящ, когато видът на разпределението е неизвестен или то се различава от нормалното и зависимата променлива (резултат) е върху ординална скала.

– Знаков тест. Предназначен е за проверка на хипотези за медианата в популацията и често включва използване на не-

зависими извадки от типа „преди-после“. Нулевата хипотеза твърди, че разликата в медианите е нула. Не изисква нормално разпределение. Използва се при рангова скала.

– Знаково-рангов тест на Wilcoxon. Използва се при зависими извадки и ординална скала на измерване на зависимата променлива. Основана се също на сравняване на медианите.

– Критерий на Kolmogorov-Smirnov. Използва се при една или две извадки за проверка на съответствието между емпиричното и теоретичното разпределение. Подходящ е при непрекъснати променливи величини, но е възможен и при дискретни разпределения. Не изисква нормално разпределение.

Размерът на ефекта от една здравна технология се отнася до действителните размери на наблюдаваните различия между групите или до силата на зависимостите между променливите.

Важно е да се знае, че макар и да е достигната статистическа значимост, ефектът от дадена здравна технология може да бъде клинично незначим или с толкова минимален ефект, че да не представлява интерес. Освен това съществуват и критерии извън статистическата значимост, които трябва да се отчитат при оценяването на здравните технологии. Те се повлияват до голяма степен от ценностите и икономическите ограничения в управлението на здравеопазването.

3. Методи за изучаване на причинни зависимости

Разделят се на два основни вида анализ – корелационен и регресионен.

1. Корелационен анализ

Една от главните задачи на статистическите методи, използвани при оценяването на здравните технологии, е разкриването и изучаването на взаимовръзките между явленията и причинната обусловеност на едни явления и процеси от други. Следователно явленията могат да се разглеждат като причини (независими променливи – x) и следствия (зависими променливи – y).

Причинно-следствените връзки между явленията също могат да се разделят на две категории – функционални и корелационни зависимости.

Функционална зависимост е тази, при която на всяко измерение на независимата променлива (x) съответства точно определено измерение на зависимата променлива (y). При оценяване на резултатите от здравните технологии на практика функционална зависимост не се среща.

Корелационна зависимост е тази, при която определено изменение в независимата променлива (x) не винаги води до точно определена промяна в зависимата променлива (y), т.е. на определено измерение на x могат да съответстват няколко измерения на y . Корелационният анализ се използва в около 10% от научните статии.

Корелационните зависимости могат да се класифицират в няколко основни групи:

а) Според формата на проявлението си корелацията може да бъде:

– Праволинейна корелация – съществува постоянна пропорционалност между абсолютните изменения на x и y . Например височината и теглото при децата, или социалния статус и смъртността при възрастните. Математически се представя чрез праволинейни управления от типа: $y = a \pm b \cdot x$

– Криволинейна корелация – при нея равномерните изменения на x водят до неравномерни изменения на y . Математически криволинейната корелация се представя чрез параболични, хиперболични, реципрочни и експоненциални криви.

б) Според начина и посоката на влияние корелацията може да бъде:

– Пряка корелация – наблюдава се, когато с увеличаване (намалвяване) на стойностите на x стойностите на y също се увеличават (намалвяват).

– Обратна корелация – наблюдава се, когато с увеличаване (намалвяване) на стойностите на x стойностите на y намалвяват (се увеличават).

в) Според начина на изследване на връзките корелацията може да бъде:

– Обикновена корелация – при нея се измерва връзката между две променливи, характеризиращи една причина (x) и едно следствие (y), без да се взема предвид влиянието на други фактори и причини.

– Частична корелация – при нея се изследва връзката между две променливи, характеризиращи една причина (x) и едно следствие (y), при условие че другите независими променливи запазват константно ниво.

– Множествена корелация – измерва връзката между една зависима променлива (y) и множество фактори, които я обуславят (x, x', x''...xⁿ). При оценяването на здравните технологии обикновено се използват множествени корелации.

Корелационната зависимост се изразява с „коэффициент на корелация“ (R). Винаги ($0 \leq R \leq 1$). При $R = 0$ липсва корелационна зависимост, а при $R = 1$ зависимостта е функционална. Междинните стойности на R изразяват силата на корелационната зависимост между x и y – слаба ($R \leq 0,3$), умерена ($0,31 \leq R \leq 0,5$), значителна ($0,51 \leq R \leq 0,7$), голяма ($0,71 \leq R \leq 0,9$) и изключително голяма ($R > 0,9$).

Корелационният анализ е приложим както за количествени, така и за качествени променливи. Методите за изчисляване на R също са многобройни:

а) Изчисляване на R при качествени алтернативни променливи. Данните се представят в четирикратна таблица. Всяка една от двете променливи – причина (x) и следствие (y), има само по две разновидности – табл. 20.

Коефициентът на корелация R при качествени алтернативни променливи се изчислява по формула на Chuprov:

$$R = \sqrt{\frac{\chi_{\text{emp}}^2}{n}}, \text{ където}$$

n – брой на всички случаи в таблицата

χ_{emp}^2 – изчислена стойност от емпиричните данни.

Таблица 20. Четирикратна таблица при качествени алтернативни променливи

X \ Y	Y		Общо
	Y ₁	Y ₂	
X ₁	Съчетание X ₁ Y ₁ (a)	Съчетание X ₁ Y ₂ (b)	a + b
X ₂	Съчетание X ₂ Y ₁ (c)	Съчетание X ₂ Y ₂ (d)	c + d
Общо	a + c	b + d	a + b + c + d

При същата ситуация може да се използва и модифицирана формула на коефициента на корелацията на Pearson:

$$R = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$$

Причинно-следствената връзка при качествени алтернативни признаци може да се изчисли и чрез друг специален критерий, представляващ съотношение на две допълващи се вероятности (OR)

$$OR = \frac{ad}{bc}$$

Показателят OR дава добра представа за настъпване на едно или друго неблагоприятно явление при експонирани и неекспонирани лица.

Като хипотетичен пример можем да разгледаме зависимост на нежеланите лекарствени реакции и метода на антибиотично лечение. Проведено е проучване върху появата на нежелани лекарствени реакции (y) при две групи пациенти (x) – едната група приема антибиотици и витамини (x₁), а другата група (x₂) – само антибиотици. Данните от проследяването на пациентите са представени в табл. 21.

Изчислената стойност на $\chi^2_{em} = 7,74$. При $df = 1$ теоретичната стойност на $\chi^2_{0,05} = 3,84$. Следователно $\chi^2_{em} > \chi^2_{0,05}$.

Таблица 21. Изследване на нежелани лекарствени реакции при антибиотично лечение

X \ Y	С нежелани лекарствени реакции	Без нежелани лекарствени реакции	Общо
Терапия с антибиотици и витамини (X ₁)	9 (a)	57 (b)	66 (a + b)
Терапия само с антибиотици (X ₂)	16 (c)	29 (d)	45 (c + d)
Общо	25 (a + c)	86 (b + d)	111 (a + b + c + d)

Това дава основание за извода, че има закономерна връзка между нежеланите лекарствени реакции и лечението.

След като чрез χ^2 е установена значителна връзка, трябва да се проверят степента и посоката на корелацията.

Чрез формулата на Chuprov:

$$R = \sqrt{\frac{\chi_{\text{ем}}^2}{n}} = \sqrt{\frac{7,74}{111}} = 0,26$$

Знакът на коефициента на корелацията се определя чрез модифицираната формула на Pearson:

$$R = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}} = \frac{9 \times 29 - 16 \times 57}{\sqrt{(16 + 29)(9 + 57)(16 + 9)(29 + 57)}} = -0,26$$

Чрез формулата OR:

$$OR = \frac{ad}{bc} = \frac{29 \times 9}{16 \times 57} = 0,29$$

Стойността на $R < 1$ показва, че лечението с антибиотици в комбинация с витамини намалява статистически значимо появата на нежелани лекарствени реакции. Стойностите на R показват, че зависимостта е обратна и слаба.

б) Изчисляване на R при качествени променливи с повече от две разновидности

Когато зависимата и независимата променлива са описателни и имат повече от две разновидности, тогава корелацията се установява чрез коефициента на Pearson (C) и чрез коефициента на Kramer (V).

$$C = \sqrt{\frac{\chi_{em}^2}{\chi_{em+n}^2}} \quad V = \sqrt{\frac{\chi_{em}^2}{n(r-1)}}, \text{ където}$$

n – броят на всички случаи

r – броят на редовете в таблицата.

Използваме хипотетичен пример за установяване на влиянието на образованието върху здравето (виж „Теория на човешкия капитал“). Проведено е изследване сред 300 лица относно нивото на образование (x) и заболяемостта (y). Независимата променлива x има три разновидности: x_1 – основно образование, x_2 – средно образование, x_3 – висше образование. Зависимата променлива y също има три разновидности: y_1 – ниска заболяемост, y_2 – средна заболяемост, и y_3 – висока заболяемост.

Резултатите от изследването са представени в табл. 22.

Таблица 22. Влияние на образованието върху здравето

X \ Y	Y_1	Y_2	Y_3	Общо
X_1	10	30	60	100
X_2	30	40	30	100
X_3	60	30	10	100
Общо	100	100	100	100

Изчислената стойност е $\chi_{em}^2 = 82,52$. Теоретичната стойност

на χ_t^2 при доверителна вероятност $P = 0,99$ и степени на свобода $dF = (3-1)(3-1) = 4$ е $\chi_t^2 = 13,3$. Следователно $\chi_{em}^2 > \chi_t^2$ и можем да направим извода, че заболяемостта зависи от нивото на образование. За да се установи степента на тази зависимост, се изчисляват коефициентите на Pearson и Kramer.

$$C = \sqrt{\frac{\chi_{em}^2}{\chi_{em}^2 + n}} = \sqrt{\frac{82,52}{82,52 + 300}} = 0,47$$

$$V = \sqrt{\frac{\chi_{em}^2}{n(r-1)}} = \sqrt{\frac{82,52}{300 \times (3-1)}} = 0,37$$

Тези стойности показват, че зависимостта на y от x е умерена и обратна – при висока степен на образованието има ниска заболяемост.

в) Коефициент на корелация на Pearson при количествени променливи величини

При оценяването на здравни технологии много често се използват количествени променливи. В тези случаи изчисляването на коефициента на корелацията (R) е значително по-сложно и в практиката се използват софтуерни продукти.

$$R = \sqrt{\frac{a \sum y + b \sum xy + n\bar{y}^2}{\sum y^2 - n\bar{y}^2}}$$

За онагледяване на принципа ще разгледаме прост хипотетичен пример, при който е възможно ръчното изчисление.

Проучването проследява зависимостта на теглото от ръста при деца на възраст от 6 мес. до 43 мес. – табл. 23.

Таблица 23. Изходни данни за изчисляване на коефициента на корелацията между теглото и ръста при деца

Изследвани лица	Ръст, cm (x)	Тегло, kg (y)	x.y	x ²	y ²
1 (6 мес.)	66,9	7,1	475,0	4475,6	50,4
2 (7 мес.)	68,5	7,2	493,2	4692,2	51,8
3 (12 мес.)	72,0	7,8	561,6	5184,0	60,8
4 (16 мес.)	77,0	8,3	639,1	5929,0	68,9
5 (18 мес.)	79,0	8,9	703,1	6241,0	79,2
6 (22 мес.)	82,1	9,2	755,3	6740,4	84,6
7 (24 мес.)	82,7	9,5	785,6	6839,3	90,2
8 (26 мес.)	84,2	10,4	875,7	7089,6	108,2
9 (30 мес.)	86,0	11,0	946,0	7396,0	121,0
10 (32 мес.)	86,5	10,8	934,2	7482,2	116,6
11 (34 мес.)	89,5	11,4	1020,3	8010,2	130,0
12 (35 мес.)	89,7	11,8	1058,5	8046,1	139,2
13 (43 мес.)	95,0	13,0	1235,0	9025,0	169,0
n = 13	$\sum x = 1059,1$	$\sum y = 126,4$	$\sum xy = 10\ 483$	$\sum x^2 = 87\ 150,8$	$\sum y^2 = 1270,1$

Първата стъпка е изчисляване на средната аритметична стойност на зависимата променлива тегло (y).

$$\bar{y} = \frac{\sum y}{n} = \frac{126,4}{13} = 9,72$$

Втората стъпка е определяне на коефициента „b”.

$$b = \frac{n\sum xy - \sum x \cdot \sum y}{n\sum x^2 - \sum x \cdot \sum x} = \frac{13 \cdot 10\ 483 - 1059,1 \cdot 126,4}{13 \cdot 87\ 150,8 - 1059,1 \cdot 1059,1} = 0,21$$

След това се определя коефициентът „a”.

$$a = \frac{\sum y - b\sum x}{n} = \frac{126,4 - 0,21 \cdot 1059,1}{13} = -7,38$$

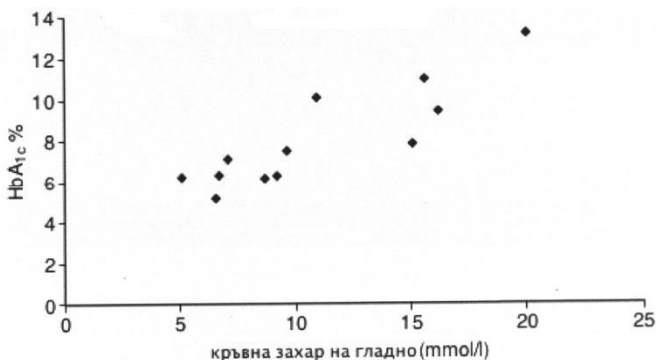
В разглеждания пример коефициентът на корелацията на Pearson при количествени променливи величини ще бъде:

$$R = \sqrt{\frac{a\sum y + b\sum xy + n\bar{y}^2}{\sum y^2 - n\bar{y}^2}} = \sqrt{\frac{7,38 \cdot 126,4 + 0,21 \cdot 10482,6 + 13,9,72^2}{1270,1 - 13,9,72^2}} = 0,98$$

Следователно при $R = +0,98$ корелационната връзка е права и изключително силна – с увеличаване на ръста се увеличава и теглото.

Друг пример за графично илюстриране на права корелационна връзка е зависимостта на нивото на кръвната захар и гликирания хемоглобин – фиг. 33. При изследването на 12 пациенти е установено $R = 0,88$, което показва висока степен на корелация.

Фигура 33. Корелационна зависимост между ниво на кръвна захар и гликиран хемоглобин



г) Рангов коефициент на корелация на Spearman

Ранговият коефициент на корелацията на Spearman (ρ) се използва при категорийни променливи, представени на ординатна ос. В основата на методиката лежи превръщане на количествените и качествените променливи в рангове (ранжиране).

$$\rho = 1 - \frac{6\sum d^2}{n(n^2 - 1)}, \text{ където}$$

n – брой изследвани случаи

d – разлика между ранговете на x и y

Винаги ранговият коефициент на Spearman е $(-1 \leq \rho \leq +1)$.

В табл. 24 са представени теоретични стойности на ρ , когато $dF = n - 2 < 8$

Таблица 24. Теоретични стойности на ρ_t при $dF = n - 2 < 8$

$dF = n - 2$	Доверителна вероятност $P = 0,95$	Доверителна вероятност $P = 0,99$
3	1,000	–
4	0,886	1,000
5	0,750	0,893
6	0,714	0,833

Когато изчислените стойности $\rho_{em} > \rho_t$, тогава корелационната зависимост е значима.

Можем да разгледаме друг хипотетичен пример в подкрепа на теорията за човешкия капитал – връзка между социално-икономическия статус и тежестта на заболяване.

В проучването са включени само 8 лица. Данните са представени в табл. 25.

Таблица 25. Връзка между социално-икономически статус и тежест на заболяване

Изследвани лица	Ранг на x	Ранг на y	$d = x - y$	d^2
1	6	5	1	1
2	7	8	-1	1
3	2	4	-2	4
4	3	3	0	0
5	5	7	-2	4
6	4	1	3	9
7	1	2	-1	1
8	8	6	2	4
Общо	-	-	-	$\sum d^2 = 24$

$$\rho_{em} = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \cdot 24}{8 \cdot (8^2 - 1)} = 0,72$$

Статистическата значимост на ρ се определя, като изчислената стойност ρ_{em} се сравнява с теоретичната стойност ρ_t .

При $dF = 8 - 2 > 6$, $\rho_t = 0,714$ при $P = 0,97$ и $\rho_t = 0,833$ при $P = 0,99$. Изчислената стойност $\rho_{em} = 0,72 > \rho_t = 0,714$ при $P = 0,95$.

Следователно между социално-икономическия статус и тежестта на заболяването има значима (при $P = 0,95$) права корелационна връзка.

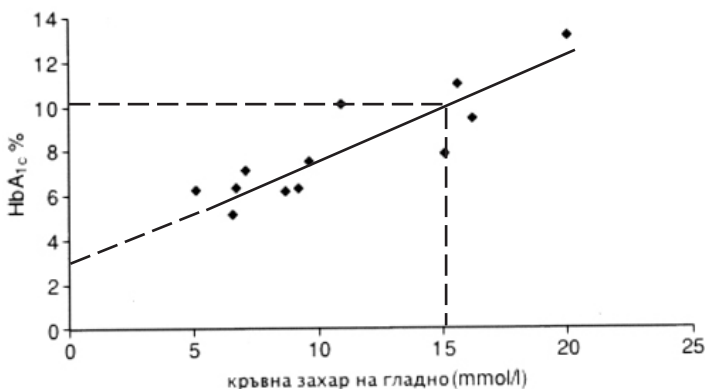
2. Регресионен анализ

Използва се за задълбочено изследване на връзката между зависимата променлива (y) и независимата променлива (x) чрез съответно моделиране. Стойностите на y се нанасят върху вертикалната ос на координатната система, а стойностите на x – върху хоризонталната ос.

Най-елементарният регресионен анализ е обикновената линейна регресия. Същността на моделирането чрез линейна регресия е да се възпроизведе права линия, която най-добре да съответства на диаграма на разсейване (права линия, при която сумата от квадратите на вертикалните разстояния от всяка точка до линията да е най-малка) или права линия, която най-добре “пасва” на всички точки от графиката.

На фиг. 34 е представена графика на примерна линейна регресия.

Фигура 34. Линия на регресионен анализ на стойностите на кръвната захар на гладно и гликирания хемоглобин (HbA_{1c}) при 12 пациенти с диабет



Линейната регресионна линия се описва с два параметъра:

- Наклон на линията (b). Измерва с колко нараства (намалява) y за всяка единица промяна на x . Нарича се още „коэффициент на регресия“.

– Начална точка на линията (а). Представява стойността на y при $x = 0$

В практиката се използват множество видове регресионни модели, които се определят от няколко критерия:

а) Според броя на включените в анализа фактори:

– Единична регресия. Изследва се връзката между едно следствие (y) и една причина (x).

$y = f(x, e)$, където:

e – компонент, отразяващ влиянието на случайните фактори върху връзката между x и y .

– Множествена регресия. Изследва се връзката между едно следствие (y) и много причини (x_i).

$y = f(x_1, x_2, x_3 \dots x_i, e)$

б) Според формата на връзката

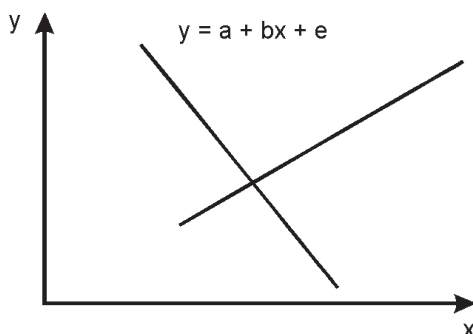
– Линейна регресия. Изразява се с просто линейно уравнение (фиг. 35).

$y = a + bx + e$

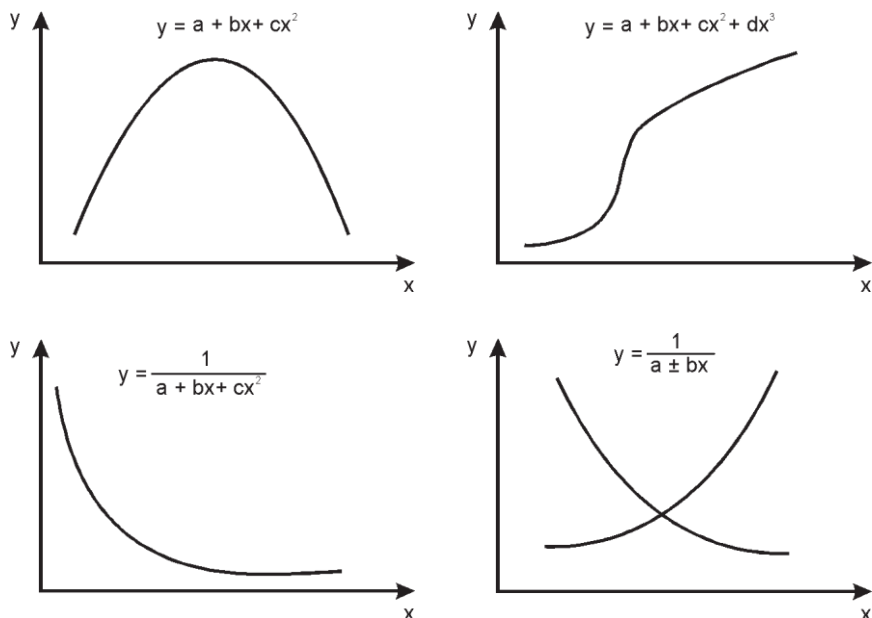
– Нелинейна регресия. Представя се с уравнение на крива. На фиг. 36 са илюстрирани някои видове нелинейни регресии и съответните уравнения на криви.

За всеки изследовател, който ще използва регресионен модел, е много важно да избере адекватен модел, който максимално се приближава до действителността. Проверката на адекватността се базира на проверката на хипотези.

Фигура 35. Линейна регресия



Фигура 36. Нелинейни регресии



Един от приложимите варианти е да се използва F-тестът (виж F-критерий на Fisher).

Друг въпрос, който трябва да се реши, когато се правят изводи от резултатите на регресионния анализ, е този за значимостта на оценките на регресионните коефициенти, тъй като съществува риск да се получат резултати, които не отразяват реалните съотношения и стойностите на параметрите могат да бъдат следствие на случайни фактори.

Значимостта на параметрите на регресионния модел се определя чрез теорията за проверка на хипотези, а методиката е чрез t-теста (виж t-критерий на Student) – отношение на стойността на параметъра „b” и репрезентативната му грешка.

Едни от най-често използваните регресионни модели в оценките на здравните технологии са множествената линейна регресия и единичната и множествената нелинейна регресия.

а) Множественa линейна регресия

Аналитичният вид на този модел се представя с уравнението:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_ix_i$$

Същността на множествената регресия е изследване на корелационната връзка между повече от две явления. Както при единичната линейна регресия, така и при множествената линейна регресия, изчисляването на параметрите се осъществява по метода на най-малките квадрати.

б) Нелинейна регресия

Съществуват различни видове нелинейни единични и множествени модели. Най-често използваните са:

– Параболична нелинейна регресия

$$y = a + b_1x + b_2x_2$$

– Хиперболична нелинейна регресия

$$y = \frac{1}{a} + b_1x + b_2x_2$$

– Експоненциална нелинейна регресия

$$y = a \cdot b_1^x$$

– Логистична нелинейна регресия

$$y = \frac{a}{1 + b_1 \cdot b_2^x}$$

– Вътрешна линейна регресия – нелинейни модели, които могат да се трансформират в линейни

$$y = e_{xp} (a + b_1x) \quad \Rightarrow \quad \lg y = a + b_1x$$

(нелинейна регресия) (линейна регресия)

Най-важното изискване за приложение на множествена регресия е независимост на променливите величини x_i .

Когато между тях съществува връзка (колинеарност), тогава регресионният анализ е неприложим.

Регресионните анализи много често се използват в оценките на здравните технологии, когато те се докладват пред комисии

с цел вземане на решения за тяхното заплащане с публични средства. В тази връзка са разработени и подробни препоръки какво трябва да се докладва за всеки регресионен анализ. Препоръките са класифицирани в пет основни групи :

1. Предстатистическа оценка за моделиране

Целта на тази оценка е да се убедят експертите, вземащи решения за реимбурсиране, че е избран подходящият регресионен модел. На този етап трябва да се отговори на следните въпроси:

– Формулирани ли са правилно целите на статистическия анализ?

– Обяснена ли е необходимостта от провеждането на нов регресионен анализ и какво е предимството пред използването на данни от вече провеждани други статистически анализи?

– Защо е използван регресионен анализ за установяване на причинните зависимости?

– Посочен ли е източникът на използваните данни? Това включва резюме на важните изследвани характеристики, като клинични, демографски, социални и др., както и метода за събиране на данните.

– Докладван ли е целият наличен размер на извадката?

– Представени ли са достатъчно аргументи за всички използвани променливи?

– Представени ли са достатъчно цифрови, графични и таблични обобщения на връзката между зависимите и независимите променливи.

– Описано ли е качеството на използваните данни? Тук се включват възможни систематични грешки, липсващи данни, невключени данни в анализа по специфични съображения и др.

– Аргументирани ли са предположенията в разработения регресионен статистически модел?

2. Статистическо моделиране

Съществуват разнообразни възможности за построяване на един регресионен модел с едни и същи данни. На този етап важни въпроси, които интересуват оценяващите модела, са:

– Предоставени ли са достатъчно подробности за използваните изчислителни методи?

– Аргументация за избора на регресионния модел. Защо например е предпочетена логистична нелинейна регресия пред множествена нелинейна регресия?

– Оправдан ли е изборът на ковариации на променливите?

– Докладване и аргументиране на действителния размер на извадката, използвана в регресионния модел.

– Аргументиране на избрания подход за липсващи данни. Когато в използваните данни има липсващи стойности, тогава тези данни могат да се изключат (доказва се, че това няма да доведе до систематична грешка при обработването на другите данни) или да се въведат (избраният метод за въвеждане трябва да бъде съобразен с чувствителността на регресионния анализ).

3. Представяне на окончателния регресионен модел

Важно е да се докладват всички подробности от окончателния статистически регресионен модел съгласно изискванията на съответната комисия, която взема решения за реимбурсиране на здравните технологии.

– Оценени ли са правилно коефициентите на регресия? За някои видове модели е по-уместно да се анализира промяната на регресионните коефициенти за различните променливи – например при логистична регресия се прилага методът на „нечетни съотношения“ на регресионните коефициенти.

– Избрани ли са подходящи мерки за неопределеност и значимост? Обикновено неопределеността се изразява с доверителни интервали, стандартни грешки или Р-стойности. Препоръчва се в оценката на здравните технологии да се включва формална оценка на неопределеността, като се доказва, че тази оценка запазва всички причинни зависимости между входящите параметри на модела. За постигане на тази цел трябва взаимовръзката вариация-ковариация да бъде представена в подходящ формат.

4. Валидиране на крайния регресионен модел

Целта на изследваните въпроси в този раздел е установяване на надеждността на докладваните резултати. Критичната оценка на регресионния модел се състои във:

– Представени ли са обобщени мерки за доброкачественост на съответствието? Изчислението на този показател зависи от

размера на извадката (или диапазона на използваните данни) и структурата на регресионния анализ (линеен или нелинеен). Например при логистична регресия се използва тест на Hosmer-Lemeshow (за подробности виж Applied Logistic Regression, Wiley, New York, 2000).

– Представени ли са резултати от валидиране на регресионния модел спрямо външни данни? Подходящ начин за валидиране е сравняване на моделираните данни от регресионния анализ с наблюдения, използващи външни данни. Всички системни разлики между моделираните и наблюдаваните резултати са индикация, че регресионният модел трябва да се коригира. Когато липсват външни данни, тогава могат да се използват алтернативни статистически подходи за валидиране – bootstrap или jackknife (за подробности виж Permutation, Parametric and Bootstrap Tests of Hypotheses, Springer, 2005).

– Дискутирана ли е достоверността на моделираните резултати и/или коефициенти? Например, ако регресионният анализ е използван, за да се моделират дългосрочни резултати от анализ разход/ползност, който използва мултиатрибутивната класификационна система EQ-5D (виж гл. V, т. 8), достоверността на резултатите може да се оцени чрез сравняването им с публикувани EQ-5D стойности за общата популация.

– Сравнени ли са резултатите от регресионния анализ с други литературни данни? Регресионните статистически анализи, чиито резултати нямат номинална валидност, обикновено съдържат грешка.

5. Потвърждаване на неопределеността в анализа

Важно е неопределеността да бъде правилно представена в регресионните статистически модели. Препоръчителните дискуссионни теми са следните:

– Докладван ли е методът за обработка на параметричната неопределеност? За параметрични коефициенти обикновено се използва методът за стандартна грешка. При смесена параметрична неопределеност се използва методът на Bayesian (за подробности виж Reference analysis, Handbook of statistics, 2005).

– Проучени ли са структурните неопределености? Оценките на здравните технологии, които са базирани на статистическо моделиране, задължително трябва да включват анализ, изследващ влиянието на различните структурни предположения. Ако се проучва структурна неопределеност с използване на сценарен анализ, тогава методите трябва да се докладват.

– Дискутирани ли са ограниченията на използвания модел? Всички ограничения на регресионния модел и неговия диапазон на приложение трябва да се докладват заедно с всеки потенциален източник на систематични грешки.

Разгледаните подробни препоръки за докладване на резултатите от статистическия регресионен анализ са важни, защото резултатите от моделирането се използват като главни параметри при оценката на здравните технологии и вземането на решения за финансиране с публични средства.

ИЗВОДИ

1. Основното приложение на статистиката в сферата на оценяване на здравните технологии е свързано с характеристиките на популацията, рисковия профил, влиянието на фактори от околната среда върху здравето, изследване на динамиката в популацията и данните за разпределение на ресурсите в здравеопазването.

2. В голяма степен обективността на статистическото заключение зависи от качеството на извадката. Най-общо всяка извадка трябва да бъде:

- Подбрана въз основа на случаен непреднамерен подбор, за да се намали вероятността от системна грешка;
- Представителна, за да се подобри валидността;
- Достатъчно голяма по обем, за да се повиши точността на изчисляваните описателни характеристики.

3. За да бъде един резултат статистически значим, не трябва да има повече от 5% вероятност той да се дължи на случайност.

4. Колкото е по-ниска стойността на p-value, толкова е по-малък шансът нулевата хипотеза да е достоверна. $P = 0,01$ означава, че вероятността за шанс е 1 на 100.

5. Важно е да се знае, че макар и да е достигната статистическа значимост, ефектът от дадена здравна технология може да бъде клинично незначим или с толкова минимален ефект, че да не представлява интерес. Освен това съществуват и критерии извън статистическата значимост, които трябва да се отчитат при оценяването на здравните технологии. Те се повлияват до голяма степен от ценностите и икономическите ограничения в управлението на здравеопазването.

Надяваме се, че сега вече знаете **отговорите** на поставените в началото въпроси:

1. Какво е приложението на статистиката в ОЗТ?
2. Кои са основните понятия в статистиката?
3. Как се проверяват хипотези?
4. Какво е p-value?
5. Какво е корелационен и регресионен анализ?