

30. Пирсонов χ^2 -тест.

Методи статистичког закључивања који су до сада били приказани су тзв. класични методи инференцијалне статистике. Код њих се полази од претпоставке да је познат облик теоријске функције расподеле F посматраног обележја X , при чему F зависи од коначног броја непознатих параметара.

Ако, пак, није познат тип расподеле обележја X примењују се поступци који се називају **непараметарски методи** или **методи независни од расподеле**. Хипотезе о расподели обележја (које се, према томе, не тичу вредности параметара, већ самог типа расподеле) које се тада постављају, називају се **непараметарске хипотезе**, а одговарајући тестови су **тестови сагласности** (стварне расподеле обележја са претпостављеном расподелом) или **непараметарски тестови**.

Постоји обимна литература која се тиче теорије и примена непараметарског тестирања. Непараметарски тестови су веома популарни. Један од разлога је то што је овај метод статистичког закључивања применљив при општијим претпоставкама од оних које захтевају параметарски методи. Други разлог је техничке природе, јер је рачуница у многим случајевима једноставнија. Са друге стране, њихов главни недостатак је што често долази до губитка информације из узорка, па се могу показати мање ефикасним од неких других стандардних техника.

На основу простог сл. узорка $(X_1, X_2, \dots, X_n)$ из обележја X , за задати праг значајности α , тестира се нулта хипотеза $H_0(F = F_0)$ против алтернативне хипотезе $H_1(F \neq F_0)$.

Пирсонов χ^2 -тест сагласности, иако један од најстаријих тестова сагласности, данас представља један од тестова са најширом облашћу примене. У статистику га је увео Карл Пирсон. Може се користити било да је обележје X дискретног или апсолутно-непрекидног типа. Задатак се тиче провере да ли је $F = F_0$, где је F_0 дата функција расподеле (она може бити једнозначно одређена или може зависити од s непознатих параметара).

Нека је са r међусобно дисјунктних подскупова $S_1, S_2, \dots, S_r \subseteq \mathbb{R}$ одређено једно разбијање скупа могућих вредности обележја X (ово разбијање је **донекле произвољно**). Нека је са M_k означен број сл. величина из простог случајног узорка $(X_1, X_2, \dots, X_n)$ које узимају вредност у скупу S_k и нека је $p_k := P\{X \in S_k | H_0\}$, $k = \overline{1, r}$. Тада важи $M_k \in \mathcal{B}(n, p_k)$, $k = \overline{1, r}$.

Тест-статистика је дата са

$$\chi_U^2 = \sum_{k=1}^r \frac{(M_k - np_k)^2}{np_k}$$

и она добро карактерише одступања случајних величина M_k од својих мат. очекивања np_k , $k = \overline{1, r}$. При томе, може се одредити њена асимптотска расподела, за довољно велики обим узорка n , под претпоставком да је хипотеза H_0 тачна. При тачној нулној хипотези H_0 , требало би разликовати следеће ситуације:

- ако је F_0 једнозначно одређена (тј. H_0 је проста хипотеза)

$$\chi_U^2 \xrightarrow{D} \chi_{r-1}^2, \quad \text{при } n \rightarrow \infty;$$

- ако F_0 зависи од $1 \leq s < +\infty$ непознатих параметара $\theta = (\theta_1, \theta_2, \dots, \theta_s)$

$$\chi_U^2 \xrightarrow{D} \chi_{r-s-1}^2, \quad \text{при } n \rightarrow \infty;$$

у овом случају вероватноће $p_k = p_k(\theta)$ зависе од θ , па и тест-статистика χ_U^2 зависи од s -димензионог параметра θ ; под извесним условима регуларности непознати параметри се прво оцењују нпр. методом максималне веродостојности из оригиналног узорка (тако се обично поступа у пракси, јер је лакше; иначе би требало параметре оцењивати на основу груписаних података).

С обзиром да је облик тест статистике

$$\chi_U^2 = \sum_{k=1}^r \frac{(\text{стварно} - \text{очекивано})^2}{\text{очекивано}},$$

велике реализоване вредности ове случајне величине указују на велику разлику стварног и очекиваног, па тада има основа за одбацавање хипотезе H_0 .

Поступак тестирања: За реализовани узорак $x = (x_1, x_2, \dots, x_n)$ прво се одређују класе S_k , при чему је емпиријско правило да број класа r буде што је могуће већи (тима се повећава увереност у расподелу F_0), али да је **очекивана вредност $np_k \geq 5$, за свако $k = \overline{1, r}$** , и да нема великих разлика у очекиваним вредностима по класама. За задати праг значајности α , из таблице квантила χ^2 -расподеле, одређује се критична вредност C тако да важи

$$P\{\chi_U^2 \geq C | H_0\} = \alpha.$$

За дати реализовани узорак x сл. величине M_k имају своје реализације m_k , $k = \overline{1, r}$, па је реализована вредност тест статистике $\sum_{k=1}^r \frac{(m_k - np_k)^2}{np_k}$. Правило одлучивања гласи: ако је реализована вредност тест статистике већа или једнака C нулту хипотезу H_0 би требало одбацити; у супротном се, на основу расположивих података и за задати праг значајности α , нулта хипотеза H_0 прихвата као тачна.

Пример. Нека је X посматрано обележје, које представља дужину трајања поправке одређеног машинског елемента. На располагању су вредности овог обележја, регистроване на узорку од 40 елемената. Тестира се хипотеза H_0 да X има Пуасонову $\mathcal{P}(3)$ расподелу, за задати праг значајности $\alpha = 0.1$. (НАПОМЕНА: у неким ситуацијама поправка елемента се може извршити на лицу места и тада се сматра да је дужина трајања поправке једнака 0 дана)

Решење.

Подаци из узорка су дати у следећој табели

дужина трајања поправке (у данима)	0	1	2	3	4	5	6	≥ 7
број елемената	1	3	7	6	10	7	6	0

Овде се, за разлику од расподела апсолутно-непрекидног типа, појављује природан избор класа, с тим што би класе које имају очекивани број мањи од 5 требало спојити. Обично су то класе на почетку, односно крају. Подаци са очекиваним вредностима и знаком како би требало извршити поделу на класе, дати су у новој табели

	0	1	2	3	4	5	6	≥ 7
m_k	1	3	7	6	10	7	6	0
p_k	0.04979	0.14936	0.22404	0.22404	0.16803	0.10082	0.05041	0.03351
np_k	1.9916	5.9744	8.9616	8.9616	6.7212	4.0328	2.0164	1.3404
	7.966				7.3896			

Реализована вредност тест-статистике је $\chi_U^2 = 9.241725$, што је значајно веће од 7.77944 – квантила реда 0.90 χ^2 -расподеле са 4 степена слободe. Хипотезу H_0 требало би одбацити. НАПОМЕНА: како је алтернативна хипотеза негација нулте, прихватање алтернативне хипотезе, на основу реализованог узорка и за задати праг значајности α , би значило да теоријска расподела обележја X **или није Пуасонова или јесте Пуасонова, али не са вредношћу параметра која је предложена у H_0** , односно оцењена из узорка.

Ако би задатак био да се тестира хипотеза H_0 да X има Пуасонову $\mathcal{P}(\mu)$ расподелу, за задати праг значајности $\alpha = 0.1$, прво би требало оценити μ , и то узорачком средином

$$\hat{\mu} = \frac{1}{40}(0 \cdot 1 + 1 \cdot 3 + 2 \cdot 7 + 3 \cdot 6 + 4 \cdot 10 + 5 \cdot 7 + 6 \cdot 6) = 3.65.$$

Одговарајућа таблица је

	0	1	2	3	4	5	6	≥ 7
m_k	1	3	7	6	10	7	6	0
p_k	0.02599	0.09487	0.17313	0.21065	0.19221	0.14032	0.08536	0.07747
np_k	1.0396	3.7948	6.9252	8.4260	7.6884	5.6128	3.4144	3.0988
	11.7596						6.5132	

Реализована вредност тест-статистике је $\chi^2_U = 1.825846$, што је мање од 6.251389 – квантила реда 0.90 χ^2 -расподеле са 3 степена слободе. Хипотезу H_0 требало би прихватити. Исти закључак може се донети рачунањем p -вредности која је једнака

$$P\{\chi^2_3 \geq 1.825846\} = 0.6093268 > 0.1 = \alpha.$$

На графику је приказан дијаграм релативних учестаности формиран на основу узорка, а назначене су и тачке које одговарају вероватноћама из $\mathcal{P}(3.65)$ расподеле. Види се да постоје извесна одступања између стварних и теоријских вредности, а критеријум помоћу кога се испитује да ли су одступања у границама случајности или нису обезбеђен је Пирсоновим χ^2 -тестом сагласности (њиме се, заправо, пореде хистограми). △

