HODNOCENÍ VÝKONNOSTI ATRIBUTIVNÍCH ZNAKŮ JAKOSTI

Josef Křepela, Jiří Michálek

OSSM při ČSJ

Červen 2009

Hodnocení způsobilosti atributivních znaků jakosti (počet neshodných jednotek)

Nechť p je pravděpodobnost výskytu neshodné jednotky v dávce. Relativní četnost je pak nejlepším bodovým odhadem této pravděpodobnosti

$$\hat{p} = \frac{z}{n},$$

z je počet zjištěných neshodných jednotek ve výběru rozsahu n. Od relativní četnosti se odvozuje konfidenční interval, pokrývající se zadanou pravděpodobností (konfidenční úrovní $1 - \alpha$) skutečnou hodnotu pravděpodobnosti p.

Vychází se z předpokladu, že proces je statisticky zvládnut a řízen např. pomocí regulačního diagramu.

S touto pravděpodobností p lze spojit jednoznačným způsobem hodnotu ukazatele výkonnosti procesu, a to

$$p = 2\Phi(-3P_p) \Leftrightarrow p/2 = \Phi(-3P_p)$$

na základě představy, že výrobek je neshodný s pravděpodobností p, když sledovaný znak jakosti – rozměr padne mimo mezní hodnoty, $\Phi(\cdot)$ je distribuční funkce normovaného normálního rozdělení N(0, 1).

Příklad

Uvažujme dávku rozsahu $n = 1365$ jednotek a počet zjištěných neshodných jednotek $z = 0, 1, 2, \ldots$.

Potom z výše uvedeného vztahu plyne

$$-3P_p = \Phi^{-1}(p/2)$$

$$P_p = -\frac{1}{3}\Phi^{-1}(p/2)$$

$$P_p = -\frac{1}{3}u_{p/2}$$

kde $u_{p/2}$ je odpovídající kvantil rozdělení N(0, 1).

Tento vztah slouží k přenesení konfidenčního intervalu pro p na konfidenční interval pro P_p .

Konfidenční interval pro parametr p binomického rozdělení při n pokusech (rozsah výběru) a z zjištěných neshodných jednotkách, při konfidenční úrovni $1 - 2\alpha$ má tvar

$$\left\langle \frac{z}{z+(n-z+1)\,F_{1-\alpha}(\nu_1\,,\nu_2)}\ ;\ \frac{(z+1)\,F_{1-\alpha}(\nu_3\,,\nu_4)}{n-z+(z+1)\,F_{1-\alpha}(\nu_3\,,\nu_4)} \right\rangle$$

kde $F_{1-\alpha}(\nu_1, \nu_2)$, resp. $F_{1-\alpha}(\nu_3, \nu_4)$ jsou $(1-\alpha)$ - kvantily F-rozdějení o stupních volnosti $\nu_1 = 2(n-z+1)$; $\nu_2 = 2z$; $\nu_3 = 2(z+1)$; $\nu_4 = 2(n-z)$.

Uvedený konfidenční interval můžeme vypočítat s využitím funkce FINV programu MS Excel.

Pro n = 1365 dostaneme:

{0,000000; 0,002699} pro z = 0,
{0,000019; 0,004075} pro z = 1,
{0,000177; 0,005283} pro z = 2,
{0,000453; 0,006409} pro z = 3,
{0,000799; 0,007486} pro z = 4,
{0,001190; 0,008527} pro z = 5.

Těmto hodnotám konfidenčních intervalů pro p odpovídají konfidenční intervaly pro ukazatele P_p následovně:

{1,0000; neexistuje} pro z = 0,
{0,9574; 1,4272} pro z = 1,
{0,9298; 1,2497} pro z = 2,
{0,9087; 1,1689} pro z = 3,
{0,8915; 1,1177} pro z = 4,
{0,8768; 1,0804} pro z = 5.

Z toho plyne, že při kontrole n = 1365 jednotek a při nenalezení žádné neshodné jednotky lze tvrdit, že výkonnost procesu není s pravděpodobností 0,975 horší nežli P_p = 1. Naopak při nalezení z = 5 neshodných jednotek lze tvrdit, že s pravděpodobností 0,025 výkonnost procesu není lepší než P_p = 1,08.

Aby tedy výkonnost procesu byla srovnatelná s požadavkem např. P_p > 1,33, bylo by nutno překontrolovat zhruba 47 000 jednotek, z nichž by žádná nesměla být neshodná.V řeči matematické statistiky to znamená zamítnout hypotézu P_p=1,33 ve prospěch alternativní hypotézy, že P_p>1,33 na hladině významnosti α=0,05.

Tento postup je v praxi těžko použitelný, ukazuje se opět, že hodnocení procesu na základě atributivních znaků je nevýhodné.

Výpočet s podporou MS Excel je proveden s využitím funkce FINV(*prst; volnost 1; volnost 2*).

Konfidenční interval na konfidenční úrovni 1-α

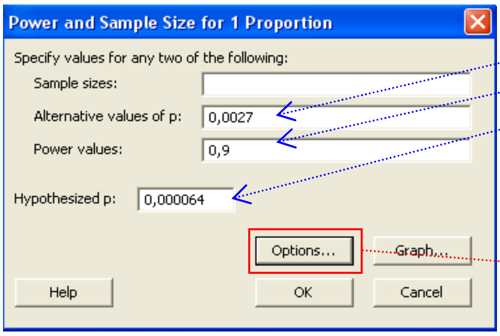
n =	**1365**
α =	**0,05**

z	$p_{dolní}$	$p_{horní}$	**P_p dolní**	**P_p horní**
0	xxx	0,002699	**1,0000**	xxx
1	0,000019	0,004075	**0,9574**	1,4272
2	0,000177	0,005283	**0,9298**	1,2497
3	0,000453	0,006409	**0,9087**	1,1689
4	0,000799	0,007486	**0,8915**	1,1177
5	0,001190	0,008527	**0,8768**	1,0804
6	0,001615	0,009543	**0,8640**	1,0511
7	0,002064	0,010537	**0,8526**	1,0269
8	0,002534	0,011515	**0,8422**	1,0064
9	0,003019	0,012479	**0,8328**	0,9886
10	0,003519	0,013431	**0,8240**	0,9728

Konstrukce přejímacího plánu pro neshodné jednotky

Máme zajistit způsobilost procesu a navrhnout test hypotézy, že podíl neshodných jednotek je přijatelný $p_1 = 0{,}000064$ (odpovídá $P_p = 1{,}33$) proti alternativní hypotéze, $p_2 = 0{,}0027$ (odpovídá $P_p = 1{,}0$) při riziku $\alpha = 0{,}05$ a síle testu $1\text{-}\beta = 0{,}90$. Zjistíme kolik je třeba odebrat a překontrolovat jednotek (rozsah výběru $n = ?$).

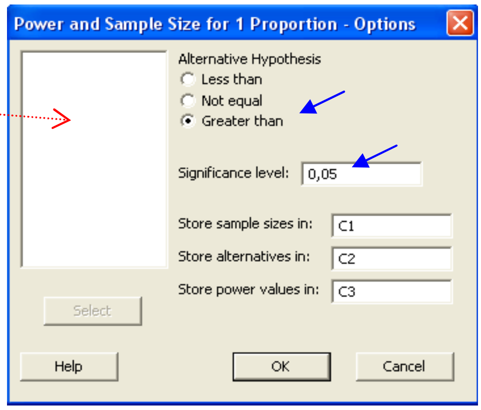
K řešení této otázky je třeba použít vhodný software, např. Minitab (nástroj „*Power and Sample Size for 1 Proportion*“). Zadání se zapíše do následujících dialogových oken:



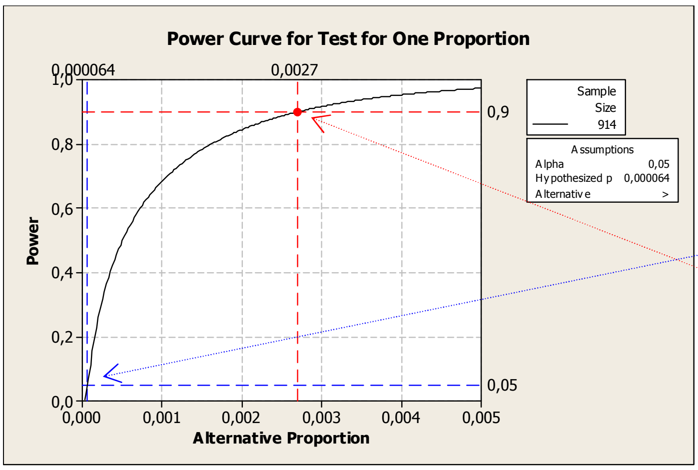
$p_2 = 0{,}0027$

$1\text{-}\beta = 0{,}90$

$p_1 = 0{,}000064$



Grafický výstup:



Silofunkce testu prochází body:
$\langle p_1 = 0{,}000064;\ \alpha = 0{,}05\rangle$ a
$\langle p_2 = 0{,}000064;\ 1\text{-}\beta = 0{,}90\rangle$

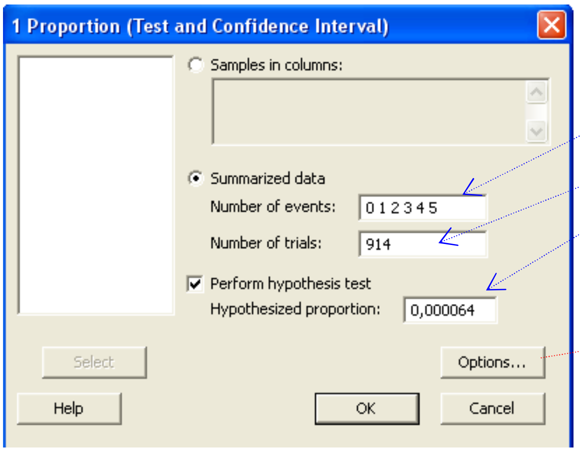
Číselný výstup:

Power and Sample Size

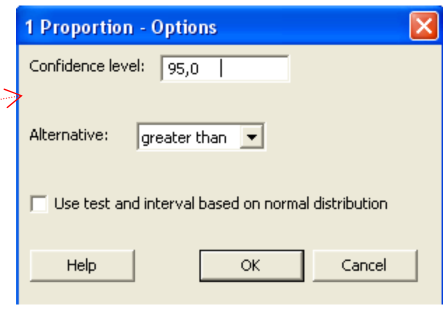
```
Test for One Proportion

Testing proportion = 0,000064 (versus > 0,000064)
Alpha = 0,05


Alternative  Sample  Target
 Proportion    Size   Power  Actual Power
     0,0027     914     0,9      0,900111
```

Zadání vyhovuje rozsah výběru $n = 914$ jednotek. Nyní uvažujme, že provedeme test a v náhodném výběru tohoto rozsahu zjistíme $z = 0, 1, 2, \ldots$ neshodných jednotek.

K vyhodnocení testu s podporou Minitabu použijeme nástroj „*Basic Statistics - 1 Proportion*“, kde vyplníme dialogová okna



Uvažované počty neshodných jednotek $z = 1, 2, 3, \ldots$;
ve výběru rozsahu $n = 914$;
testovaný podíl $p_1 = 0{,}000064$.



Číselný výstup:

Test and CI for One Proportion

Test of p = 0,000064 vs p > 0,000064

Sample	X	N	Sample p	95% Lower Bound	Exact P-Value
1	0	914	0,000000	*	1,000
2	1	914	0,001094	0,000056	0,057
3	2	914	0,002188	0,000389	0,002
4	3	914	0,003282	0,000895	0,000
5	4	914	0,004376	0,001496	0,000
6	5	914	0,005470	0,002158	0,000

Vyhodnocení testu:

V případě, že bude ve výběru $n = 914$ jednotek zjištěno $z = 0$, případně $z = 1$ neshodných jednotek (z je značeno ve výpisu X), není důvod zamítnout testovanou hypotézu že $p_1 = 0{,}000064$, jelikož odpovídající p-hodnoty jsou větší, než zvolená hladina významnosti $\alpha = 0{,}05$.

V případě nalezení dvou či více neshodných jednotek ($z = 2, 3, \ldots$) se zamítá testovaná hypotéza $p_1 = 0{,}000064$ ve prospěch alternativní hypotézy $p_2 = 0{,}0027$. Jinými slovy, v tomto případě se zamítá hypotéza že $P_p \geq 1{,}33$ ve prospěch hypotézy $P_p \leq 1{,}0$.

Hodnocení způsobilosti atributivních znaků jakosti (počet neshod)

Tak jako byl analyzován problém neshodných jednotek, tak lze postupovat i v případě neshod, když na jedné jednotce je možno zjisti více neshod (vad, defektů). V tomto případě jsou výpočty založeny na Poissonově rozdělení pravděpodobnosti s parametrem λ, který představuje střední hodnotu výskytu neshod. Nejlepším bodovým odhadem parametru λ je pak zjištěný počet neshod, v kontrolovaném výběru (podskupině), případně průměrný počet zjištěných neshod ve výběru, při m kontrolovaných výběrech. Konfidenční interval, pokrývající s pravděpodobností $1 - 2\alpha$ skutečnou hodnotu parametru λ má tvar

$$\left\langle \frac{1}{2m}\chi^2_{\alpha}(\nu_1) \ ; \ \frac{1}{2m}\chi^2_{1-\alpha}(\nu_2) \right\rangle$$

kde $(1 - 2\alpha)$ je zvolená konfidenční úroveň; $\nu_1 = 2mx_0$ a $\nu_2 = 2(mx_0 + 1)$ jsou stupně volnost; x_0 je průměrný počet zjištěných neshod v dávce, z m kontrolovaných dávek. $\chi^2_{\alpha}(\nu_1)$ a $\chi^2_{1-\alpha}(\nu_2)$ jsou α a $1-\alpha$ kvantily χ^2 – rozdělení o ν_1 a ν_2 stupních volnosti. Tyto kvantily a tedy celý konfidenční interval můžeme vypočítat s využitím funkce CHIINV programu MS Excel.

Budeme interpretovat λ jako $n \cdot p$, kde p je pravděpodobnost výskytu neshody a n je maximální možný počet neshod. V případě m kontrolovaných dávek označíme k maximální možný počet neshod v dávce. Potom $n = m \cdot k$. V konkrétním případě bude potenciální počet neshod n fixován a měnit se bude parametr $\lambda = n \cdot p$. To se přímo promítne do konfidenčních mezí ukazatele výkonnosti procesu

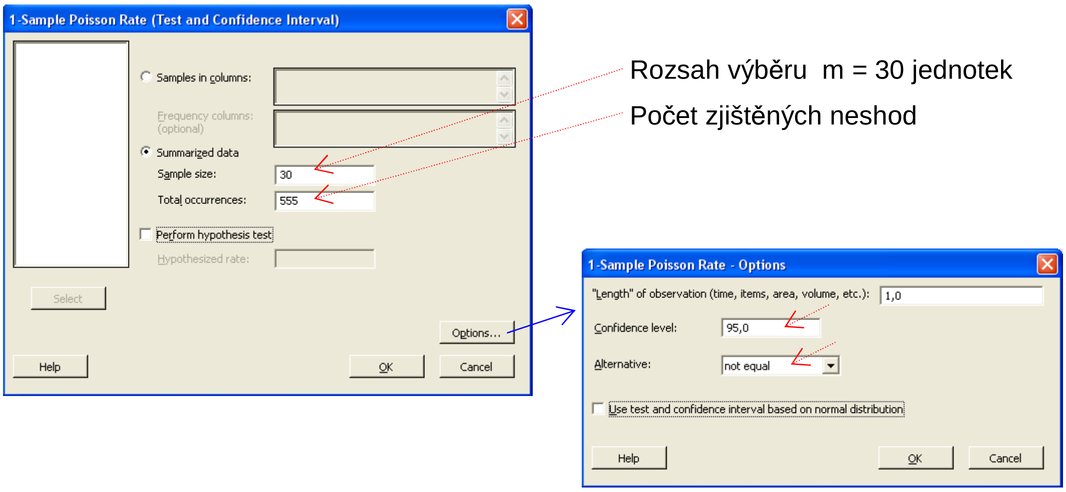
$$p = 2\Phi(-3P_p) .$$

Příklad

Uvažujme kontrolu $m = 30$ jednotek (dávek), které obsahují po stejném počtu $k = 100$ součástek (tištěné spoje – desky do televizorů, osazené součástkami). Bylo zjištěno celkem $z = 555$ neshod. To znamená, že odhad očekávaného počtu neshod na jednu jednotku je $555 / 30 = 18{,}5$. Pro konfidenční úroveň $1-2\alpha = 0{,}95$ $(\alpha = 0{,}025)$ a pro $\nu_1 = 2mx_0 = 60 \cdot 18{,}5 = 1110$, $\nu_2 = 2(mx_0 + 1) = 1110 + 2 = 1112$, je $\chi^2_{\alpha}(\nu_1) = \chi^2_{0{,}025}(1110) = 1019{,}561$ a $\chi^2_{1-\alpha}(\nu_2) = \chi^2_{0{,}975}(1112) = 1206{,}310$. Potom konfidenční interval pro parametr λ je

$$\left\langle \frac{1}{2m}\chi^2_{\alpha}(\nu_1) = 16{,}9927 \ ; \ \frac{1}{2m}\chi^2_{1-\alpha}(\nu_2) = 20{,}1052 \right\rangle .$$

Výpočet konfidenčního intervalu s podporou softwaru Minitab provedeme pomocí nástroje „*1 – Sample Poisson Rate*“ kde vyplníme dialogová okna následovně:



Číselný výstup:

```
Confidence Interval for One-Sample Poisson Rate

                  Total          Rate of
Sample  Occurrences   N  Occurrence         95% CI
     1          555  30     18,5000  (16,9927, 20,1052)


"Length" of observation = 1.
```

Jelikož celkový možný počet neshod je $n{\cdot}m = 100{\cdot}30 = 3000$, potom máme při $\lambda = n{\cdot}m{\cdot}p$ pro odhad pravděpodobnosti výskytu jedné neshody

$$\hat{p} = \frac{\hat{\lambda}}{n \cdot m} = \frac{18{,}5}{3000} = 0{,}00617\,.$$

Při použití konfidenčních mezí pro λ získáme interval, ve kterém leží pravděpodobnost p při úrovní konfidence $1 - 2\alpha = 0{,}95$:

$$16{,}9927 / 3000 = 0{,}00566 < p < 20{,}1052 / 3000 = 0{,}00670.$$

Odtud již snadno pomocí vztahu

$$P_p = -\frac{1}{3}\Phi^{-1}\left(\frac{\lambda}{2nm}\right)$$

získáme konfidenční interval pro ukazatel výkonnosti P_p :

$$-\frac{1}{3}\Phi^{-1}\left(\frac{0{,}00566}{2}\right) > P_p > -\frac{1}{3}\Phi^{-1}\left(\frac{0{,}00670}{2}\right)$$

$$0{,}9038 < P_p < 0{,}9215\,.$$

Výkonnost procesu není uspokojivá, neboť ukazatel P_p je s pravděpodobností 0,975 pouze lepší než 0,9038 a požadováno je, aby byl roven alespoň 1. Výrobu je nutno zlepšit. Maximální počet případných neshod by musel být

$$P_p > 1 \quad \Leftrightarrow \quad -\frac{1}{3}\Phi^{-1}(x) = 1,$$

tedy $x = \Phi(-3)$ tj. $x = -0{,}00135$. Pak pravděpodobnost výskytu neshody nesmí být větší než 0,0027. Tomu odpovídá při celkovém počtu možných neshod 3000 očekávaný počet neshod na jednotku

$$\lambda = 3000 \cdot 0{,}0027 = \mathbf{8{,}1}.$$

Z toho plyne, že přijatelný počet neshod zjištěných při kontrole 30 jednotek by se měl pohybovat okolo 243 neshod, aby bylo možno říci, že proces pracuje s výkonností nejméně $P_p = 1$.

S podporou Minitabu, nástroje „*1 – Sample Poisson Rate*“, můžeme tento závěr upřesnit a vypočítat dolní 95% - ní mez konfidenčního intervalu pro zvolené počty zjištěných neshod (např. 255, 260, až 280). Zjistíme, že při výskytu méně než 270 neshod, při kontrole 30 jednotek, bychom neměli důvod pochybovat, že ukazatel výkonnosti $P_p = 1$, na hladině významnosti $\alpha \approx 0{,}05$.

Test and CI for One-Sample Poisson Rate

Test of rate = 8,1 vs rate > 8,1

Sample	Total Occurrences	N	Rate of Occurrence	95% Lower Bound	Exact P-Value
1	255	30	8,50000	7,64382	0,229
2	260	30	8,66667	7,80194	0,145
3	265	30	8,83333	7,96014	0,085
4	270	30	9,00000	8,11842	0,046
5	275	30	9,16667	8,27678	0,023
6	280	30	9,33333	8,43522	0,011

"Length" of observation = 1.