Exploratorní analýza výběrového souboru dat pevnosti drátkobetonu v tlaku

Ing. Daniel PIESZKA
Ing. Ivan KOLOŠ, Ph.D.
doc. Ing. Karel KUBEČKA, Ph.D.
VŠB-TU Ostrava – Fakulta stavební

Věrohodné vyhodnocení experimentálních dat a následné stanovení statistických odhadů charakteristických hodnot materiálových vlastností je při hodnocení existujících konstrukcí z hlediska funkční způsobilosti zásadní. Článek se zabývá možnou metodikou ověření předpokladů o těchto datech, zejména metodikou ověření nezávislosti prvků, stejné pravděpodobnosti zařazení prvků do výběru, stejného rozdělení hustoty pravděpodobnosti a homogenity výběrového souboru experimentálních dat jako nutných podmínek spolehlivého vyhodnocení.

Exploratory analysis of SFRC compressive strength of a sample data file

The evaluation reliability of experimental data files and the estimation of characteristic material property values are the key points in the assessment of the serviceability of existing constructions. The paper deals with possible procedures for the data assumption verification, especially procedures for the verification of data independence, the same data probability, the same data probable density and sample data file homogeneity as the required conditions for reliable data evaluation.

Úvod

Důvodem pro vlastní hodnocení existující konstrukce je podle ČSN ISO 13822 [1] hledisko současného stavu a jejího budoucího použití, tedy hledisko požadavků budoucí funkční způsobilosti. Tato způsobilost je definována úrovní bezpečnosti uživatelů při užívání konstrukce, úrovní trvale udržitelných vlastností a úrovní požadavků na použitelnost, životnost a trvanlivost konstrukce.

Funkční způsobilost konstrukce se ověřuje na modelech podle ČSN EN 1990 [2], které spolehlivě reprezentují zatížení a chování konstrukce a únosnost jejích jednotlivých prvků. Tyto výpočetní modely musí reprezentovat také změny ve způsobu budoucího užívání, pokud k němu dojde.

Předpokladem možnosti stanovení materiálových vlastností je znalost fyzikálních, chemických i biologických vlivů prostředí. Vlastnosti materiálů se tedy stanovují experimentálně, destruktivními či nedestruktivními zkouškami. V případě destruktivního zkoušení betonu v konstrukcích se pak vychází z ČSN EN 12504-1 [3], alternativně z ČSN EN 13971 [4].

Vyhodnocení zkoušek a stanovení odhadu charakteristických hodnot vlastností materiálů se podle ČSN EN 1990 [2] provádí statistickými metodami a pravděpodobnostním počtem [8]. Takto získané výsledky jsou však závislé na charakteristikách zkoumaného výběrového souboru dat. Pro korektní statistické vyhodnocení je tedy nezbytné ověřit, zda je výběrový soubor reprezentativní (tzn. zda jsou prvky výběru vzájemně nezávislé, stejně pravděpodobné, zda pocházejí ze stejného rozdělení hustoty pravděpodobnosti a zda je celý výběr homogenní).

Článek předkládá možný postup ověření vlastností výběrového souboru se stanovením předpokladů pro následné statistické zpracování a vyhodnocení. V jednotlivých krocích je prezentován na příkladu ověření souboru dat destruktivních měření pevnosti drátkobetonu v tlaku dle ČSN EN 12390-3 [5] vzorků odebraných z drátkobetonové průmyslové podlahové konstrukce v souladu ČSN EN 12504-1 [3]. Hodnoty jednotlivých měření jsou jako výběrový soubor dat kvantitativní proměnné uvedeny v *tab. 1*.

Tab. 1. Pevnost betonu v tlaku $f_{c,i}$ dle metodiky [5]

Pevnost betonu v tlaku $f_{c,i}$ [MPa] pro i-tý vzorek, kde $i = \langle 1, 50 \rangle$ (po řádcích)							
29,9	31,6	38,0	22,9	41,1	27,6	34,9	27,4
37,7	28,4	26,9	31,7	29,9	36,1	31,3	32,8
31,9	30,0	24,1	38,1	25,7	32,2	29,1	33,9
35,8	34,4	32,9	24,1	31,2	35,1	29,7	38,2
25,5	28,5	35,7	36,3	26,3	31,9	27,8	25,9
22,5	39,3	23,0	34,1	32,4	29,7	33,9	35,7
38,9	26,2	–	–	–	–	–	–

Ověření náhodnosti výběru

Hodnocená pravoúhlá plošná konstrukce byla před odběrem vzorků položena do roviny x - y kartézského souřadnicového systému a počátek byl ztotožněn s hraničním rohovým bodem. Poloha jednotlivých vývrtů byla popsána dvojicí souřadnic x_i a y_i, jejichž velikost byla získána z tabulky náhodných čísel daného intervalu v souladu s ustanovením čl. 81 ČSN 010250 [7]. Prvky výběru je tedy možno považovat za náhodné.

Tento akademický postup stanovení polohy vývrtů se v praxi nejspíše neuplatní, neboť by došlo k odběru vzorků i z těch částí konstrukce, které není žádoucí poškodit (např. koridory pojížděné manipulační technikou).

Ověření nezávislosti prvků výběru

Korelace vlastností prvků výběru bývá způsobena zejména nestabilitou měřicího zařízení a zanedbáním okrajových podmínek (teploty, času), obecně pak systémovými chybami měření. Výběrový soubor dat závislých prvků nelze následně považovat za vydatný. K ověření nezávislosti sousedních prvků jednorozměrného souboru dat použijme autokorelační test významnosti autokorelačního koeficientu 1. řádu ρ_1. Význam korelačního koeficientu vyplývá ze vztahu $x_i = \rho_1 x_{i-1} + e_i$, kde e_i je ryze náhodná složka čistě náhodného průběhu.

Formulujme následující hypotézy [9]:

– nulová hypotéza H_0, prvky výběru jsou vzájemně nezávislé, $\rho_1 = 0$;
– alternativní hypotéza H_A, prvky výběru jsou autokorelovány a korelace je významná, $\rho_1 \neq 0$.

Testovací statistika

$$t_n = \frac{T_1\sqrt{n+1}}{\sqrt{1-T_1}}\,, \tag{1}$$

kde

$$T_1 = \left(1-\frac{T}{2}\right)\sqrt{\frac{n^2-1}{n^2-4}}\,, \tag{2}$$

přičemž T je von Neumannův poměr

$$T = \frac{\sum_{i=1}^{n-1}\left(x_{i+1}-x_i\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2} \tag{3}$$

a kritickým oborem pro test autokorelace I. řádu

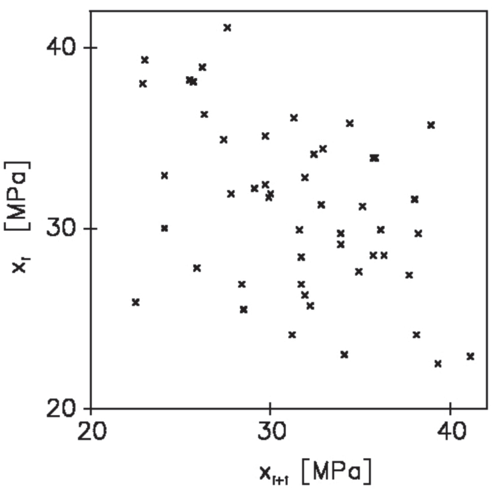
$$\left|t_n\right| > t_{1-\alpha/2}\left(n+1\right)\,, \tag{4}$$

kde α je hladina významnosti, zde α = 0,025;

po dosazení $\left|t_n\right| = 2{,}151 < 2{,}310 = t_{0{,}975}\ (51).$

Testovací statistika nepadá do kritického oboru hodnot a na hladině významnosti α = 0,025 není důvod nulovou hypotézu zamítnout. Prvky výběru nejsou autokorelované, jsou nezávislé.

Graficky je také možné posoudit autokorelaci (*obr. 1*). Z grafu je zřejmé, že zkoumané prvky výběru nevykazují žádný významný trend 1. řádu a že nejsou korelované.



Obr. 1. Graf autokorelace

Ověření homogenity výběru

Homogenitu výběru obecně narušují takové hodnoty kvantitativní proměnné, které se od ostatních hodnot mimořádně liší. Tyto mimořádné hodnoty označme jako odlehlá pozorování a k jejich identifikaci použijme následující pravidlo. Za odlehlé pozorování budeme považovat takovou hodnotu x_i, jejíž *z-score* (*tab. 2*) je větší než |3|, tedy je-li tato hodnota x_i vzdálena od výběrového průměru o více než trojnásobek výběrové směrodatné odchylky (5), (6).

$$z\text{-}score_i = \frac{x_i - \bar{x}}{s}\,, \tag{5}$$

pokud tedy platí $\left|x_i - \bar{x}\right| \geq 3s$, (6)

pak x_i je odlehlým pozorováním.

Výběrový průměr je definován jako

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{7}$$

a po dosazení $\bar{x} = \bar{f}_c = 31{,}36$ MPa .

Výběrová směrodatná odchylka je definována jako

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\,, \tag{8}$$

po dosazení pak s = 4,80 MPa .

Tab. 2. Z-score prvků výběru

Z-score$_i$ [-] pro *i*-tý vzorek, kde $i = \langle 1, 50\rangle$ (po řádcích)							
–0,31	0,05	1,38	–1,76	2,03	–0,78	0,74	–0,83
1,32	–0,62	–0,93	0,07	–0,31	0,99	–0,01	0,30
0,11	–0,28	–1,51	1,40	–1,18	0,17	–0,47	0,53
0,92	0,63	0,32	–1,51	–0,03	0,78	–0,35	1,42
–1,22	–0,60	0,90	1,03	–1,06	0,11	–0,74	–1,14
–1,85	1,65	–1,74	0,57	0,22	–0,35	0,53	0,90
1,57	–1,08	–	–	–	–	–	–

Jak je patrné z *tab. 2*, u žádného prvku výběru nedosahuje *z-score* kritické hodnoty |3|. Ve výběru tedy nejsou odlehlá pozorování a výběr je možné považovat za homogenní. Maximální hodnoty dosahuje *z-score* u hraničních prvků výběru, tj.

– pro $x_{min} = x_{41}$ = 22,5 MPa, je *z-score* rovna –1,85,
– pro $x_{max} = x_5$ = 41,1 MPa, je *z-score* rovna 2,03 (hodnoty x_{min} a x_{max} viz *tab. 1*).

Ověření normality výběru – test dobré shody χ^2

Ověřme hypotézu o předpokladu normálního rozdělení výběru, tj. předpoklad, že výběr pochází z rozdělení $N(\mu, \sigma^2)$. Vzhledem k absenci apriorní znalosti střední hodnoty μ a směrodatné odchylky σ základního souboru nahraďme tyto parametry výběrovým průměrem $\overline{x}$ a výběrovou směrodatnou odchylkou s. Nulová hypotéza H_0 – náhodný výběr pochází ze základního souboru s normálním rozdělením, alternativní hypotéza H_A – náhodný výběr nepochází ze základního souboru s normálním rozdělením.

Výběrový soubor o rozsahu n rozdělme do k třídních intervalů J_1 až J_k, kde velikost intervalu volme mezi s/4 a s/2. Dále stanovme třídní četnosti n_k a středy tříd c_k (*tab. 3*). Horní hranice intervalů x_k převeďme na hodnoty normované proměnné

$$u_k = \frac{x_k - \mu}{\sigma}, \tag{9}$$

kde neznámé parametry rozdělení základního souboru nahraďme parametry výběru stanovenými podle (7), (8), tedy

$$u_k = \frac{x_k - \overline{x}}{s}. \tag{10}$$

Tab. 3. Třídy, třídní četnosti a distribuční funkce výběrového souboru dat

Třídy [MPa]	Třídní četnost n_k [-]	Střed třídy c_k [MPa]	Hodnoty distribuční funkce $F(f_c)$ [-] k horní hranici třídy
$J_1 = \langle 22{,}0; 24{,}0\rangle$	3	23,0	0,060
$J_2 = (24{,}0; 26{,}0\rangle$	5	25,0	0,160
$J_3 = (26{,}0; 28{,}0\rangle$	6	27,0	0,280
$J_4 = (28{,}0; 30{,}0\rangle$	8	29,0	0,440
$J_5 = (30{,}0; 32{,}0\rangle$	6	31,0	0,560
$J_6 = (32{,}0; 34{,}0\rangle$	7	33,0	0,700
$J_7 = (34{,}0; 36{,}0\rangle$	6	35,0	0,820
$J_8 = (36{,}0; 38{,}0\rangle$	5	37,0	0,920
$J_9 = (38{,}0; 40{,}0\rangle$	3	39,0	0,980
$J_{10} = (40{,}0; 42{,}0\rangle$	1	41,0	1,000
	$\Sigma n_i = 50$		

Dále stanovme odpovídající distribuční funkci normovaného normálního rozdělení $N(0, 1)$

$$\Phi(u_k), \tag{11}$$

relativní třídní četnost

$$\pi_{0\,k} = \Phi(u_k) - \Phi(u_{k-1}), \tag{12}$$

absolutní třídní četnost

$$n\pi_{0\,k}; \tag{13}$$

podmínkou dalšího postupu je ověření, zda platí $n\pi_{0,k} > 5$. Pokud podmínka není splněna, příslušné intervaly sloučíme. V tomto případě tedy sloučíme intervaly J_1 a J_2 a intervaly J_8, J_9, J_{10} (*tab. 4*).

Nyní stanovme hodnotu testované statistiky

$$G = \chi^2 = \sum_{k=1}^{k} \frac{(n_k - n\pi_{0\,k})^2}{n\pi_{0\,k}}; \tag{14}$$

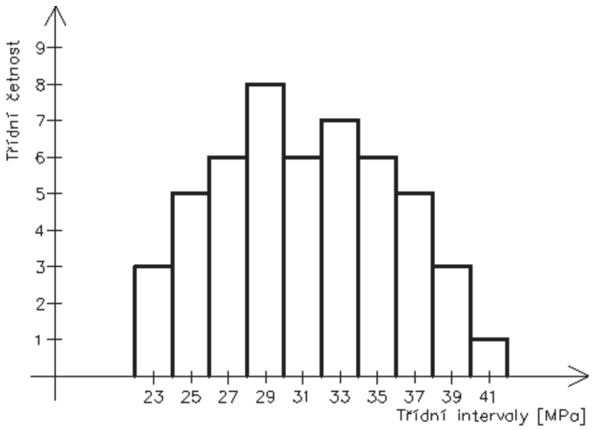
přes redukovaný počet tříd je tedy hodnota testované statistiky

$$G = \chi^2 = 1{,}292.$$

Kritickým oborem pro test normality je

$$\chi^2 > \chi_{1-\alpha}(k - h - 1), \tag{15}$$

kde α je hladina významnosti, zde $\alpha = 0{,}025$, h je počet odhadovaných parametrů (μ, σ^2), tj. $h = 2$.



Obr. 2. Histogram (tab. 3)

Tab. 4. Výpočet statistiky χ^2

Horní mez intervalu x_k [MPa]	Třídní četnost n_k [-]	Norm. horní hranice u_k	Norm. distribuční funkce $\Phi(u_k)$	Relativní třídní četnost $\pi_{0,k}$	Absolutní třídní četnost $n\pi_{0,k}$	Upravená ab. četnost $n\pi_{0,k}$ ($n\pi_{0,k} > 5$)	Upravená třídní četnost n_k	Statistika χ^2 (G_k)
24,0	3	-1,53	0,0630	0,0630	3,150	–	–	–
26,0	5	-1,12	0,1314	0,0684	3,420	6,570	8	0,311
28,0	6	-0,70	0,2420	0,1106	5,530	5,530	6	0,040
30,0	8	-0,28	0,3897	0,1477	7,385	7,385	8	0,051
32,0	6	0,13	0,5517	0,1620	8,100	8,100	6	0,544
34,0	7	0,55	0,7088	0,1571	7,855	7,855	7	0,093
36,0	6	0,97	0,8340	0,1252	6,260	6,260	6	0,011
38,0	5	1,38	0,9162	0,0822	4,110	–	–	–
40,0	3	1,80	0,9641	0,0479	2,395	–	–	–
42,0	1	2,22	0,9868	0,0227	1,135	7,640	9	0,242
							$\Sigma\chi^2$	**1,292**

Dle statistických tabulek [7] je kritická hodnota pro čtyři stupně volnosti

$$\chi_{1-0,025}\,(7-2-1)=\chi_{0,975}\,(4)=11{,}143.$$

Vypočtená hodnota testované statistiky nespadá do oboru kritických hodnot

$$\chi^2=1{,}292<11{,}143=\chi_{0,975}\,(4).$$

Není je tedy důvod na hladině významnosti α = 0,05 zamítnout nulovou hypotézu, že výběr pochází ze základního souboru s normálním rozdělením. Zamítáme hypotézu alternativní.

Pro posouzení symetrie rozložení hodnot výběru kolem výběrového průměru stanovme výběrovou šikmost dle vztahu

$$\alpha=\frac{n}{(n-1)(n-2)}\frac{1}{s^3}\sum_{i=1}^{n}(x_i-\bar{x})^3\ ; \qquad (16)$$

po dosazení α = –0,02. Hodnoty výběru jsou kolem výběrového průměru rozloženy symetricky ($\alpha \to 0$).

Pro posouzení koncentrace hodnot výběru kolem výběrového průměru stanovíme výběrovou špičatost dle vztahu

$$\beta=\frac{n(n+1)}{(n-1)(n-2)(n-3)}\frac{1}{s^4}\sum_{i=1}^{n}(x_i-\bar{x})^4 - 3\frac{(n-1)^2}{(n-2)(n-3)} \qquad (17)$$

po dosazení β = – 0,84.

Koncentrace hodnot kolem výběrového průměru neodpovídá přímo normálnímu rozdělení, pro které platí β = 0. Křivka hustoty rozdělení pravděpodobnosti výběrového souboru je plošší než u normového normálního rozdělení. Podle histogramu na *obr. 2* se však plochost nejeví jako významná.

Stanovme výběrový variační koeficient

$$V_\chi=\frac{s}{\bar{x}}=\frac{4{,}80}{31{,}36}=0{,}15. \qquad (18)$$

Míra variability proměnné x_i, tedy měřené pevnosti betonu v tlaku $f_{c,i}$, je 15 %.

Statistická analýza dat a jejich vyhodnocení

Výběrový soubor můžeme nyní považovat za reprezentativní výběr ze základního souboru s normálním rozdělením, neboť bylo ověřeno, že jednotlivé prvky výběru jsou vzájemně nezávislé, stejně pravděpodobné, pocházejí ze stejného normálního rozdělení hustoty pravděpodobnosti a výběr je homogenní.

Teprve nyní je možné přistoupit ke stanovení bodových odhadů parametrů základního souboru. I zde ovšem platí určitá pravidla. Dobrý (věrohodný) bodový odhad musí být zejména nestranný, vydatný, konzistentní a dostatečný. Každý bodový odhad parametru je sám o sobě náhodnou veličinou, neboť je stanoven z výběrového souboru dat. Tato vypočtená hodnota parametru se bude od skutečného parametru základního souboru lišit. Velikost chyby při stanovení parametru z jednoho výběrového souboru je výběrovou chybou. Je-li bodový odhad parametru nezkreslený, pak měřítkem přesnosti je směrodatná odchylka, v této souvislosti označovaná jako střední chyba odhadu.

Bodový odhad střední hodnoty

Požadované vlastnosti dobrého bodového odhadu střední hodnoty μ základního souboru splňuje výběrový průměr $\bar{x}$.

$$\mu=\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=31{,}36\ \text{MPa}, \qquad (19)$$

střední chyba odhadu s = 4,80 MPa.

Bodový odhad rozptylu

Požadované vlastnosti dobrého bodového odhadu rozptylu σ^2 základního souboru splňuje výběrový rozptyl

$$\sigma^2=s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2=23{,}08. \qquad (20)$$

Bodový odhad charakteristické pevnosti

Charakteristická pevnost betonu v tlaku je dle požadavků ČSN EN 1990 ed. 2 [2] definována jako 5% kvantil, tj.

$$f_{ck}=f_{c;0,05}=x_{0,05}\ . \qquad (21)$$

Metodika určení tohoto kvantilu je rovněž stanovena v [2]. Předepsanou metodou je metoda předpovědní, vycházející z výběrové směrodatné odchylky. Obecně je pak hodnota kvantilu definována jako

$$x_{p,\,predp}=\bar{x}-s\,t_p(\alpha,\,p,\,v)\left(1+\frac{1}{n}\right)^{1/2}, \qquad (22)$$

kde $\bar{x}$ je výběrový průměr, s je výběrová směrodatná odchylka, t_p p-procentní kvantil Studentova t-rozdělení, α je výběrová šikmost, v je počet stupňů volnosti, kde $v = n - 1$.
Dle doporučení normy se výběrová šikmost zanedbává. Vztah pro stanovení 5% kvantilu má tedy tvar

$$x_{0,05,predp}=\bar{x}-st_p\,(0;\,0{,}05;\,49)\left(1+\frac{1}{50}\right)^{1/2}.$$

Jednostranná hodnota 5% kvantilu Studentova t-rozdělení pro 49 stupňů volnosti je dle tabulek [2]

$$t_{0,05}\,(49)=1{,}6766;$$

po dosazení

$$x_{0,05,predp}=31{,}36-4{,}80\times1{,}6766)\left(1+\frac{1}{50}\right)^{1/2}=23{,}23\ \text{MPa};$$

charakteristická pevnost betonu v tlaku je pro testovanou konstrukci

$$f_{ck}=23{,}23\ \text{MPa}.$$

Závěr

Na uvedeném příkladu byly předvedeny možné způsoby ověření vlastností experimentálních dat a jejich způsobilosti pro následné statistické vyhodnocení. Předložený postup je použitelný pro jednorozměrná kvantitativní data. Při posouzení vzájemné závislosti prvků byl v článku aplikován test autokorelace 1. řádu. Posouzení závislosti prvků autokorelačními testy vyšších řádů ponechává kolektiv autorů na čtenáři.

Přínos statistických metod pro stavební praxi nespočívá jen v ověřování vlastností materiálů, ale také ve vlastním návrhu stavebních konstrukcí. Při navrhování podle mezních stavů se uplatňují ve výpočtu prostřednictvím dílčích součinitelů, které zohledňují nejistoty modelů zatížení, nejistoty modelů účinků zatížení, nejistoty způsobené nepříznivými odchylkami vlastností materiálů od jejich charakteristických hodnot, dále pak nejistoty modelů odolnosti a v neposlední řadě zohledňují nejistoty geometrických rozměrů konstrukce.

Literatura

[1] ČSN ISO 13822 Zásady navrhování konstrukcí – Hodnocení existujících konstrukcí. ČNI, 2005.
[2] ČSN EN 1990 ed. 2 Eurokód: Zásady navrhování konstrukcí. ÚNMZ, 2011.
[3] ČSN EN 12504-1 Zkoušení betonu v konstrukcích – Část 1: Vývrty – Odběr, vyšetření a zkoušení v tlaku. ÚNMZ, 2009.
[4] ČSN EN 13791 Posuzování pevnosti betonu v tlaku v konstrukcích a v prefabrikovaných dílcích. ČNMZ, 2007.
[5] ČSN EN 12390-3 Zkoušení ztvrdlého betonu – Část 3: Pevnost v tlaku zkušebních těles. ÚNMZ, 2012.
[6] ČSN 01 0250, Zb. Statistické metody v průmyslové praxi. Všeobecné základy. ÚNMZ, 1978.
[7] Linda, B. – Kubanová, J.: Statistické tabulky a vzorce. Univerzita Pardubice, 2000.
[8] Tichý, M.: Co s tou pravděpodobností? Stavební obzor, **21**, 2012, č. 10, s. 291-293. ISSN 1805-2576 (Online)
[9] Kubečka, K.: Využití statistických metod při statickém navrhování a posuzování železobetonových konstrukcí. VŠB-TU Ostrava, 2004.