

Vzorová zkoušková písemka – MA3, část Statistika

1. Pevnost dřevěných trámů (MPa) se dá zjišťovat nedestruktivně tak, že se měří Youngův modul pružnosti (MPa). Existují dvě metody, kterými se může měřit. V následující tabulce jsou údaje o výsledcích experimentu. V prvním a druhém sloupci jsou údaje o modulu pružnosti získané metodou I a metodou II, ve třetím skutečná pevnost trámu.

Metoda I	Metoda II	Pevnost
13806	9462	40,04
14508	10782	39,89
15389	12214	52,49
13875	10157	41,50
13638	12783	43,08
13026	11520	49,37
15778	13459	44,12
15543	12707	53,89
16748	14947	50,90
15493	12990	55,41
21971	18197	77,22
14695	11375	55,85
16296	13225	51,93
14323	11650	47,04
16953	14037	39,54
20618	16184	75,81

Spočtete následující základní popisné charakteristiky polohy pro obě dvě metody: medián, dolní kvartil, horní kvartil. Řekněte, jak se tyto charakteristiky pro obě metody liší.

Jsou v datech týkajících se první a druhé metody odlehlá pozorování? Pokud ano, udejte která.

Předpokládejme, že pevnost nového trámu je možno odhadnout použitím lineární regrese, přičemž nezávislou proměnnou je Youngův modul pružnosti. Pokud bychom se pro určení pevnosti pomocí lineární regrese měli rozhodnout mezi modelem, kde modul pružnosti je zjišťován metodou I a metodou II, který byste vybrali? (Rozhodněte na základě koeficientu determinace.)

2. Předpokládejme, že trámy byly vyrobeny z kmenů určitého typu o pevně daných rozměrech. Jejich pevnosti můžeme považovat za výběr z velikého množství trámů (tvořícího základní soubor), jejichž pevnost se řídí normálním rozdělením. Vyjádřeno řečí matematické statistiky, změřené pevnosti mohou být považovány za náhodný výběr z normálního rozdělení. Za tohoto předpokladu odhadněte, kolik procent trámů v základním souboru má pevnost nižší než 30 MPa. Pokud bychom ze základního souboru těchto trámů vybrali náhodně tři, s jakou pravděpodobností by všechny tři měly pevnost nižší než 30 MPa?

Řešení příkladu 1.

	metoda I	metoda II
medián	15441	12745
dolní kvartil	14099	11447,5
horní kvartil	16522	13748

Všechny spočtené charakteristiky jsou pro metodu I vyšší než pro metodu II. Znamená to, že obě metody neměří stejně (rozdíl v mediánech 2696, v dolních kvartilech 2651,5, v horních kvartilech 2774).

Kdybychom uvažovali rozdíly v obou změřených údajích pro jednotlivé trámy, pak je medián těchto rozdílů roven 2993,5, dolní kvartil 2411 a horní kvartil 3722.

Abychom zjistili, zda jsou v datech odlehlá pozorování musíme spočítat mezikvartilová rozpětí, která charakterizují variabilitu obou výběrů. Mezikvartilové rozpětí zjištěného modulu pružnosti první metodou je 2423 (MPa) a druhou metodou 2300,5 (MPa). Mezivartilové rozpětí vynásobíme 1,5 a odečteme od dolního kvartilu a přičteme k hornímu kvartilu. Tím získáme dolní a horní meze, vně kterých leží odlehlé hodnoty. Pro metodu I je dolní mez rovna 10464,5 (MPa) a horní mez rovna 20156,5 (MPa). Proto dvě měření leží nad horní mezí. Tyto dvě odlehlá pozorování mají hodnoty 20618 (MPa) a 21971 (MPa). Pro metodu II je dolní mez rovna 7996,75 a horní mez rovna 17198,75 (MPa). Vně intervalu leží jedno měření s hodnotou 18197 (MPa).

Lineární regrese lépe vystihuje závislost mezi dvěma veličinami, jestliže koeficient determinace je vyšší. Koeficient determinace je kvadrát výběrového korelačního koeficientu. (Někdy se též vyjadřuje v procentech.) Pro závislost mezi pevností trámu a změřeným modulem pružnosti metodou I je koeficient determinace roven 0,674(67,4%), zatímco pro metodu II 0,578(57,8%). Zdá se tedy, že pro určování pevnosti trámu jsou hodnoty modulu pružnosti zjišťované metodou I trochu lepší.

Řešení příkladu 2.

Za předpokladu, že jsou změřené pevnosti výběrem ze základního souboru (populace) řídícího se normálním rozdělením, můžeme odhadnout střední hodnotu (průměrnou hodnotu) základního souboru pomocí průměru $\bar{x} = 51,13$ a směrodatnou odchylku pomocí $\sigma_{n-1} = 11,40$. (Někdy se směrodatná odchylka odhaduje pomocí $\sigma_n = 11,04$.) Podíl trámů, které mají pevnost nižší než 30 MPa, odpovídá pravděpodobnosti, s jakou při náhodném vybírání vybereme trám s nižší pevností než 30 MPa, což se stane s pravděpodobností $\Phi\left(\frac{30-51,13}{11,40}\right) = 1 - \Phi(1,854) = 0,032$ (3,2%). Jestliže bychom vybírali z obrovského množství trámů tři, pak by pravděpodobnost, že všechny tři budou mít pevnost nižší než 30 MPa byla rovna $0,032^3 = 0,00003$ (0,003%).