

Testy normality

Mnoho statistických metod (např. Studentův t-test) předpokládá, že základní soubor má normální rozdělení. Není-li tento předpoklad splněn, nelze danou metodu použít. K určení, zda lze rozdělení dat považovat za normální, slouží **testy normality** (např. Shapirův-Wilkův test). Většina statistického softwaru implementuje nějakou formu testů normality.

*Je třeba si uvědomit, že zkoumané soubory často normální rozdělení **nemají**. Například schopnosti jsou v populaci rozloženy normálně, ale známky na vysvědčení jsou zpravidla vychýlené směrem k lepšímu hodnocení.*

Grafické metody

Jednoduchý způsob, jak alespoň přibližně odhadnout, zda data mají normální rozdělení, je sestavení **histogramu**. O něco přesnější je použití **Q-Q grafu** (*kvantil-kvantil*) nebo **P-P grafu** (*pravděpodobnost-pravděpodobnost*). Q-Q graf je vhodnější pro testování normality na krajích rozdělení, zatímco P-P graf více zdůrazňuje odchylky od normálního rozdělení poblíž střední hodnoty.

Histogram

Histogram je graf, který se sestrojí tak, že na vodorovnou osu nanášíme **hodnoty sledované veličiny** a na svislou osu **jejich četnosti**. Je-li zkoumaná veličina spojitá (tzn. nabývá-li nekonečně mnoha hodnot), rozdělíme osu na intervaly a četnosti určíme jako počty pozorovaných hodnot v těchto intervalech. Optimální počet intervalů lze určit podle tzv. Sturgesova pravidla. Při normálním rozdělení souboru by měl histogram připomínat Gaussovu křivku.

Q-Q graf

Princip této metody spočívá v tom, že na jednu osu nanášíme **kvantily hypotetického normálního rozdělení** a na druhou osu **kvantily zkoumaného souboru**. V případě normálního rozdělení leží všechny body grafu na přímce.

P-P graf

Postup při sestrování P-P grafu je podobný jako u Q-Q grafu. Na jednu osu nanášíme **hodnotu kumulativní distribuce hypotetického normálního rozdělení** a na druhou osu **hodnotu kumulativní distribuce zkoumaného souboru**. Opět v případě normálního rozdělení budou body ležet na přímce.

Ověření normality výpočtem

Existuje řada testů, které se liší silou a náročností provedení. Patří mezi ně např. **Shapirův-Wilkův**, **Andersonův-Darlingův**, **Kolmogorovův-Smirnovův**, **Lillieforsův** a další. Test se obvykle neprovádí ručně, ale kvůli velké náročnosti se výpočty provádějí na počítači. Příklad výpočtu v programu *R* (testovaný soubor je v proměnné *x*):

```
> shapiro.test(x)
```

```
Shapiro-Wilk normality test
```

```
data: x
```

```
W = 0.9685, p-value = 0.8762
```

Je-li p-hodnota větší než 0,05 normalita se **nezamítá**.

Hrubý odhad

Chceme-li posoudit normalitu rozdělení pouze orientačně, můžeme porovnat aritmetický průměr s mediánem. podle některých autorů^[zdroj?] by se neměly lišit více než o 10 %.

Odkazy

Související články

- Normální rozdělení
- Statistická inference
- Testování statistických hypotéz
- Studentův t-test

Použitá literatura

- NETOLICKÁ, Veronika. *Testy normality*. Přírodovědecká fakulta Univerzity Palackého v Olomouci : Katedra matematické analýzy a aplikací matematiky, 2008,
- KASAL, Pavel a Marie HLADÍKOVÁ. Koutek pro statistiky amatéry. *Pelikán : akademický bulletin 2. LF UK* [online]. 1995, roč. 3, no. 3, s. -, dostupné také z <<http://stary.lf2.cuni.cz/projekty/pelikan/peli0395/statis1.htm>>.

Citováno z „https://www.wikiskripta.eu/index.php?title=Testy_normality&oldid=378843“