***Otázky k tématu 5 – korelace*** [1]

1. Který z následujících korelačních koeficientů ukazuje na nejtěsnější (nejsilnější) vztah?

a) 0,55 b) 0,09 c) -0,77 d) 0,1 e) 1,05

2. U kterého z výše uvedených koeficientů platí, že jedinec s nadprůměrnými hodnotami X bude mít pravděpodobně i nadprůměrné hodnoty Y?

3. Pěti reprezentativním vzorkům lidí ve věku 15, 20, 30, 45 a 60 let jsme dali dotazník na měření politické konzervativnosti. Těmto 5 vzorkům v uvedeném počadí vyšly následující průměrné hodnoty konzervativnosti: 60, 85, 80, 70, 65. Korelace mezi věkem a politickou konzervativností je

a) 1.0 b) -1.0 c) lineární d) nelineární

4. U této otázky vyberte z uvedených scatterů ten, který odpovídá popisu

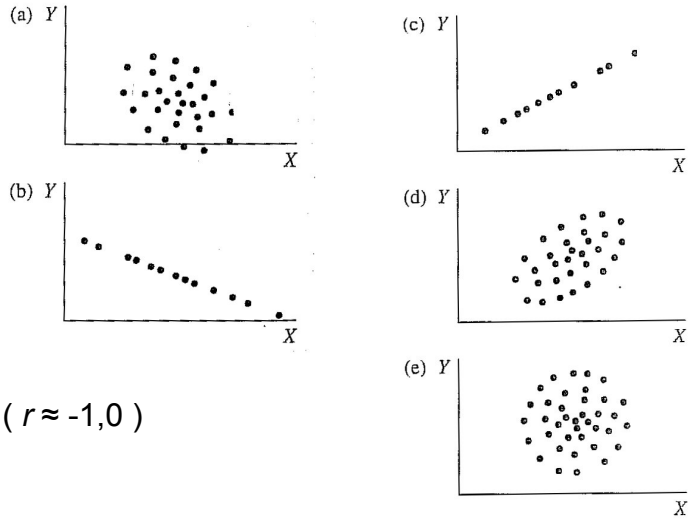
A) perfektní pozitivní lineární vztah ( $r = 1,0$ )

B) středně těsný pozitivní lineární vztah ( $r \approx 0,5$ )

C) žádný lineární vztah ( $r \approx 0$ )

D) středně těsný negativní lineární vztah ( $r \approx -0,5$ )

E) perfektní negativní lineární vztah ( $r \approx -1,0$ )




9. Odhadněte, jaká je korelace mezi níže uvedenými dvojicemi proměnných – pozitivní, negativní, nebo nulová?

a) výška v cm, váha v kg
b) věk v měsících, čas v běhu na 50 metrů
c) známka z matematiky, známka ze čtení
e) známka z matematiky, počet zameškaných hodin za rok
f) IQ, rodné číslo
g) zájem o sport, zájem o politiku
h) počet km na tachometru auta, rok výroby auta
i) maximální denní teplota, množství vody spotřebované za den domácnostmi

10. Pokud by ve skutečnosti byla odpověď na variantu h) předchozí otázky -0,8, jak by se korelace změnila, kdybychom místo proměnné „rok výroby auta“ použili proměnnou „stáří auta“?

11. Korelace mezi X a Y je 0,60; korelace mezi X a W je -0,80. Má X těsnější lineární vztah s Y nebo s W?

[1] Některá čísla otázek chybí. Není to chyba.

12. Pokud korelace mezi X a Y je 0,5, jak se změní hodnota *r*, pokud transformujeme X na *T*-skóry?

13. Pokud r=1 a $z_x$ =-0,5, kolik je $z_y$?

14. Pokud r=-1 a $z_x$ = 0,8, kolik je $z_y$?

15. Inteligenční kvocienty (IQ) získané testem A jsou konzistentně o 10 bodů vyšší než IQ získaná testem B. Jaká je teoreticky nejvyšší možná korelace mezi testem A a testem B?

16. Pan Brown spočítal kovarianci mezi výškou běžce ve stopách a časem na 800m v sekundách a na padesátihlavém vzorku studentů mu vyšlo 2,30. Na Brownových datech spočítal pan Smith kovarianci mezi výškou běžce v palcích a časem na 800m v minutách a vyšlo mu 0,46. Porovnejte Brownovu korelaci mezi výškou a časem na 800m se Smithovým výsledkem.

17. Jsou-li směrodatné odchylky dvou korelovaných proměnných $s_X = 3$ a $s_Y = 15$, jaká je maximální možná kovariance těchto proměnných? ($r_{XY} = c_{XY}/s_X s_Y$).

18. Předpokládejme, že bychom udělali odhad výšky a váhy každého studenta psychologie, který si letos zapsal PSY117. Potom bychom spočítali $r_{OVáhaOVýška}$. Jaká by byla vypočtená korelace v porovnání s korelací spočítanou na výškách a váhách změřených metrem a váhou?

19. Jedna studie o infarktech uvádí, že lidem, kteří chodí pravidelně do kostela, hrozí nižší riziko infarktu, než lidem, kteří do kostela nechodí. Který z následujících výroků je pravdivý?
a) Když začnete chodit pravidelně do kostela, vaše riziko, že dostanete infarkt se určitě sníží.
b) Mezi těmito proměnnými určitě není žádný kauzální vztah.
c) Pokud pravidelně chodíte do kostela, je méně pravděpodobné, že dostanete infarkt, než kdybyste do kostela nechodil(a).
d) Tato korelace jednoznačně ukazuje na kauzální vztah.

22. Jedna studie udává poměrně nízkou korelaci mezi IQ a kreativitou (*r* = 0,2). Směrodatná odchylka IQ skórů v jejich vzorku je však pouze 5. Co by se stalo s korelací, kdyby variabilita IQ skórů ve vzorku nebyla takto omezená?

23. Studie 280 studentů učitelství udává téměř nulovou korelaci (r = 0,1) mezi studijními výsledky (průměr známek) a schopností vyučovat (hodnocení zkušeným učitelem při cvičné hodině). Studie dále udává, že korelace mezi hodnoceními těchto 280 studentů dvěma nezávislými zkušenými učiteli je 0,21. Jak tato druhá informace ovlivní vaši interpretaci korelace mezi studijními výsledky a schopností vyučovat?

P1. a) Spočítejte na níže uvedených deseti párech pearsonův korelační koeficient *r* mezi IQ skóry a skóry z aritmetiky (nebo to nechte spočítat počítač ☺).

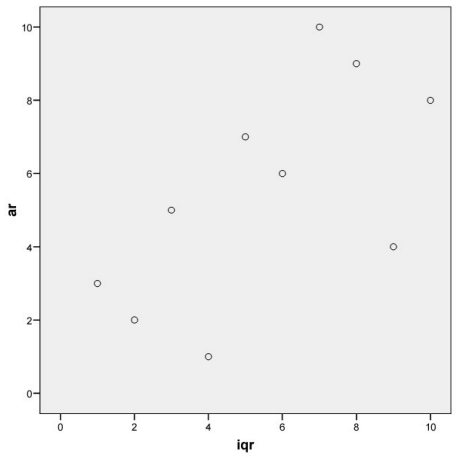
| *Pupil* | *IQ* | *Arithmetic* | *Pupil* | *IQ* | *Arithmetic* |
|---|---|---|---|---|---|
| A | 105 | 15 | F | 96 | 10 |
| B | 120 | 23 | G | 107 | 4 |
| C | 83 | 11 | H | 117 | 30 |
| D | 137 | 22 | I | 108 | 18 |
| E | 114 | 17 | J | 130 | 14 |

b) Převeďte IQ a Aritmetiku na pořadové skóry (pořadí) a načrtněte bodový diagram (scatter). Udělejte ho tak, aby jednotky na ose x i y byly stejné. Vypadá vztah lineárně? Vypadá hodnota *r*, kterou jste spočítali v bodě a), rozumně?
c) Jak se výběrový průměr a směrodatná odchylka IQ mohou promítnout do korelace, vezmeme-li v potaz, že populační průměr IQ je 100 a smodch je 15?

**Odpovědi**

*Korelace*

1. c
2. c
3. d
4. cdeab

9. a) pozitivní, b) nulová, c) pozitivní, d) nulová, e) negativní, f) nulová, g) nulová, h) negativní, i) pozitivní.
10. 0,8
11. W
12. 0,5
13. $z_Y$=-0,5
14. $z_Y$=-0,8
15. 1,0
16. Lineárními transformacemi (přičítání, násobení konstantou) nejsou ovlivněna relativní pořadí skórů ani relativní rozdíly mezi skóry, a tedy ani korelace. Velikost korelace zde tedy na jednotkách nezáleží. Oběma badatelům vyšla nakonec stejná korelace.
17. Maximální korelace může být 1, tedy max($c_{XY}$)=$s_X s_Y$, tedy 45.
18. Korelace na naměřených datech bude vyšší, protože data budou zatížena menší chybou měření.
19. c
22. Korelace by byla podstatně větší.
23. Reliabilita kritéria (tedy přesnost hodnocení) je mizerná. Zdá se, že každý hodnotitel hodnotí něco jiného. Korelace mezi výsledky a schopností učit tedy bude ve skutečnosti zřejmě vyšší.

P1. a) $r$ = 0,52  b) viz scatter, ano, ano c) $m_{iq}$=112 a to je o dost víc než v populaci, $s_{iq}$=15,7 a to je jen nepatrně víc než populaci. Korelace by tím nemela být ovlivněna.