

1. Populacja dla statystyka i metodologa:

- a. dla statystyka: zbiór danych
- b. dla metodologa: podmioty

2. Nie badamy populację, tylko próbę w celu poznania populacji.

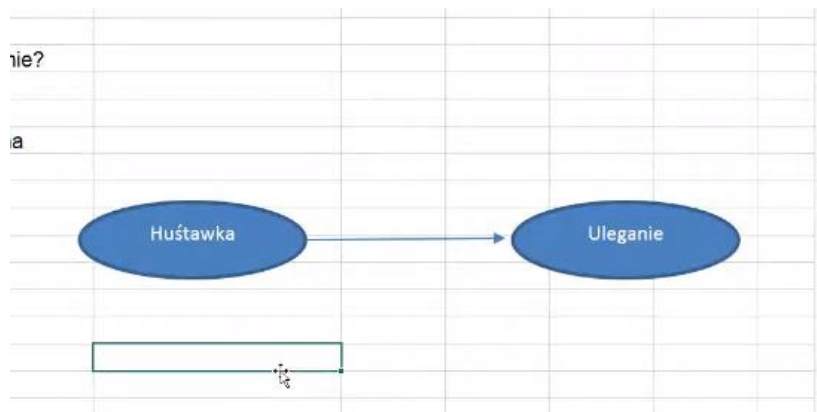
Jasne operacyjne definiowanie pojęć (konceptualizacja) jest kluczem do sukcesu.

Zawsze stawiamy pytanie badawcze, ale nie zawsze jesteśmy w stanie postawić hipotez. Konstruowanie hipotez na siłę jest bez sensu. Muszą być one empirycznie weryfikowalne. Proces indukcyjny – czy hipoteza faktycznie jest odpowiedzią na zadane pytanie badawcze. Badania dają nam tylko pewne prawdopodobieństwo, nie likwiduje wątpliwości jaka powinna być odpowiedź.

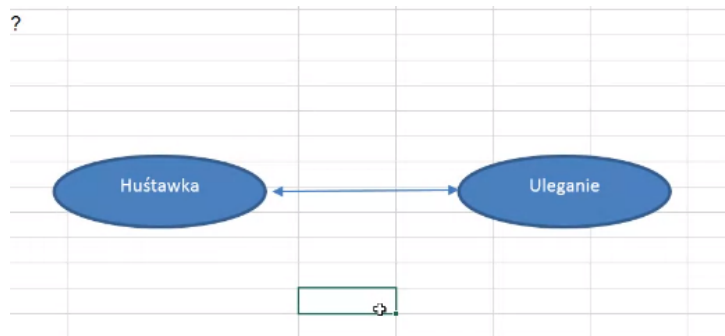
Działania psychometryczne vs działania w celu weryfikacji hipotez – są to dwa rodzaje badań.

3. Dobry i zły policjant. EKSPERYMENT

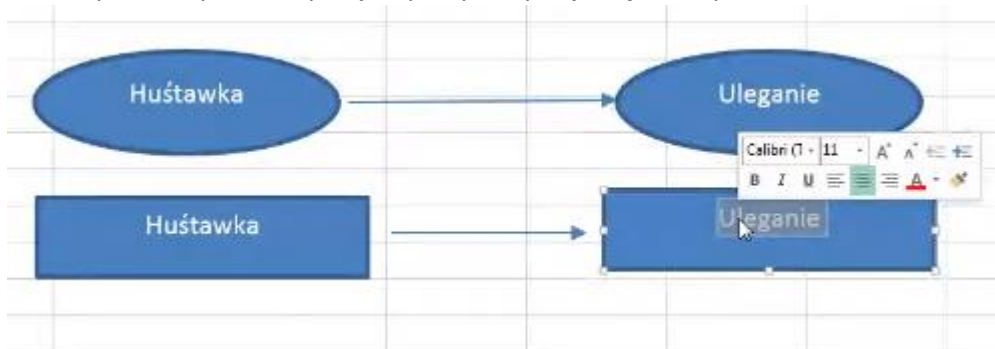
- a. Pyt. 1: Efekt tła
Pyt. 2: Huśtawka emocjonalna
- b. Pytanie: Czy huśtawka emocjonalna zwiększa uleganie?
Hipoteza: Huśtawka emocjonalna zwiększa uleganie
Zmienna 1: huśtawka emocjonalna (zmienna niezależna)
Zmienna 2: uleganie (zmienna zależna)



- c. Manipulacja – pomiar zmiennej zależnej
Jest huśtawka: uleganie
Brak huśtawki: uleganie
- d. Model współzależności. Trzeba brać pod uwagę taką możliwość.



4. Zmienne teoretyczne (latentne) oznaczamy elipsą. Prostokąty to wskaźniki, ale czasem też do zmiennych, którymi manipulujemy. Wykorzystuje się ten zapis w modelach strukturalnych



Te „same badania” (zawierające te same zmienne) mogą przynieść różne rezultaty, bo zawierają różną operacjonalizację.

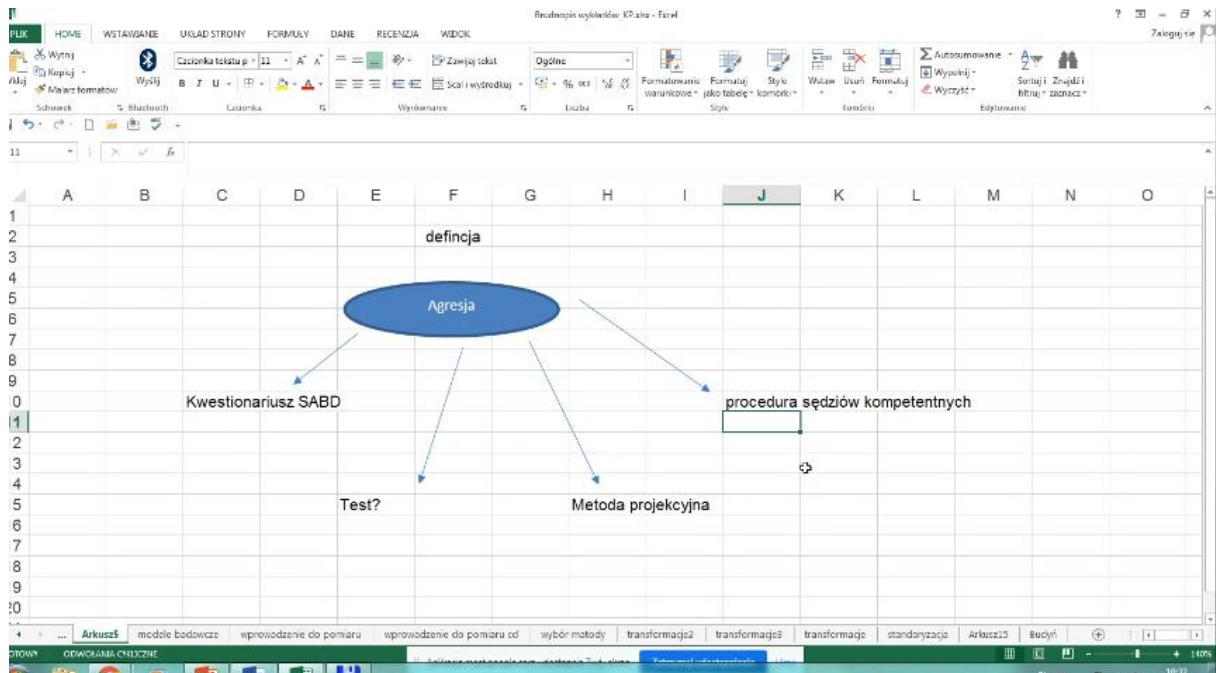
Te dwa schematy nie są takie same.

Badania Dolińskiego o huśtawce emocjonalnej

Podstawy metodologii i statystyki 2 22.02.2020

Które z pytań nie można traktować jako dobre pytania badawcze?

1. Jakim poziomem postawy twórczej charakteryzuje się badana populacja? – zdefiniowanie i zmierzenie danego konstruktu może być trudne. Jakim poziomem, ale w stosunku do czego, brakuje odniesienia, brak dwóch zmiennych
2. Co dotychczas wiadomo o chorobach nowotworowych i osobach na nie chorujących? – Pytanie z czapy, masz badać coś empirycznie a nie szukać byle czego. Nie mamy dwóch zmiennych nawet
3. Czy adekwatność obrazu własnego ciała ma wpływ na efektywność odchudzania? – trochę za mało szczegółowe jest, powinno narzucać już sposób analizy – czy efektywność odchudzania jest większa u osób które mają bardziej pozytywny obraz własnego ciała niż u osób które bla bla
4. Czy motywacja autonomiczna jest mediatorem wpływu własnej skuteczności na intencję podejmowania zachowań zdrowotnych?
5. nie można w pytaniu badawczym umieszczać badanej grupie



ROLA ZERA ABSOLUTNEGO W SKALACH POMIAROWYCH!

The spreadsheet illustrates the role of absolute zero in measurement scales. It includes a diagram showing "cukier" (sugar) and "smak budyniu" (taste of pudding) with arrows pointing to "poziom cukru" (sugar level) and "Skala Subiektywnej Oceny Budyniu (SSOB Januszek & Januszek)" respectively. The table below shows experimental results for two versions of a study.

3. Operjonalizacja		4. Paradygmat badawczy - FKSP-RYMFI		Skala Subiektywnej Oceny Budyniu (SSOB Januszek & Januszek)		skala interwałowa 1 - 10	
wersja 1		wersja 2		wynik badania	test stat.	wnioski w1	wnioski w2
grupa E	n = 50	grupa E	n = 50	M = 1,3	test stat.	brak wniosków	nie należy siodzić
grupa K	n = 50	grupa K	n = 50	M = 4,2		brak wniosków	siodzenie nie ma znaczenia
jako błąd?		jako błąd?					

Ćwiczenia 2

Rezygnuj z skali stosunkowej wieku tylko i wyłącznie gdy respondednt może stracić swoją anonimowość

Socjo demo jest potrzebne do wykazania reprezentatywności próby.

Zbieraj dane wg kryterium GUSq

17.04 Wykład Grabowski Model korelacyjny i regresja

Z reguły gdy mówimy o korelacji w badaniach mamy na myśli korelację Pearsona, natomiast mówienie, że korelacja to korelacja Pearsona to jest to duże nadużycie.

Modele korelacyjne:

Pearsona, Spearmana, dla zmiennych nominalnych

Zmienna interwałowa w psychologii. Podział skal według Stevensa. Korelacja dotyczy WSZYSTKICH rodzajów zmiennych.

Schematy badawcze w psychologii: eksperymentalny,

Towarzystwo łączy się z satysfakcją z życia, satysfakcją z pracy – zmienną pomijaną w tym przykładzie jest liczba przyjaciół, bo to oni decydują o satysfakcji z życia. I wiąże się z towarzyskością. Mamy zatem do czynienia z mediatorem zmiennej pośredniczącej, budującej relację zmiennej A z B przez C.

Mediacja ma wiele motywów analitycznych; jednym ze schematów (patrz podręcznik Szalmesi szalmesi rozdział 4; edit: oczywiście miało być shaughnessy&shaughnessy :D)

Różnice indywidualne „wyrósł” na modelu korelacyjnym

Wzrost etyki pracy = niższe wypalenie zawodowe

Im większe przywiązanie afektywne do organizacji tym mniejsze wypalenie zawodowe

Zależności można przedstawić na wykresach rozrzutu. Związek można przedstawić procentowo, podnosząc współczynnik korelacji do kwadratu (r^2 to tzw. Współczynnik determinacji, który wyraża wspólny procent wariancji dwóch zmiennych).

Cechy osobowości w relacji z wymiarami temperamentalnymi PTS (PTS odnosi się do cech układu nerwowego).

Im większa neurotyczność tym niższa siła procesu hamowania, tym niższa wydajność pracy układu nerwowego i niższa ruchliwość procesów nerwowych. Neurotyczność jest emocjonalnością, jest to cecha oznaczająca łatwość reagowania za pomocą emocji negatywnych. Wyższa skłonność do reagowania w ten sposób oznacza szereg właściwości układu nerwowego czyli niższą siłę układu nerwowego. Im większa siła tym więcej mogą znieść.

Korelacja to inaczej współwystępowanie. Wszyscy słyszeliśmy narzekania palaczy, że gdy tylko zapalą na przystanku, zaraz przyjeżdża autobus. Widać wyraźnie, że w tym wypadku mamy do czynienia ze skalą pomiarową nominalną.

Wiele korelacji to współwystępowanie zmiennych mierzonych na skalach ilościowych takich jak waga i wzrost.

Większość zjawisk w otaczającym nas świecie występuje w różnorodnych związkach.

Statystyczny opis umożliwia lepsze ich zrozumienie i modyfikowanie. Często słyszymy stwierdzenie „Rak płuc jest powiązany z paleniem papierosów”. Oznacza to, że im więcej papierosów się pali, tym bardziej prawdopodobne jest zachorowanie na raka. Mówimy, że im więcej jednego, tym więcej drugiego.

Zamiast używać nieprecyzyjnych słów (więcej, mało itp.), statystycy wolą w ocenie używać liczb, dlatego powstała matematyczna teoria korelacji i regresji, stanowiąca narzędzie dokładnego określania stopnia powiązania zmiennych ze sobą.

Podstawowym problemem statystyki jest stwierdzenie czy między zmiennymi zachodzi jakiś związek i czy jest on bardziej czy mniej ścisły. Analiza regresji i korelacji to jedna z najważniejszych i najszerzej stosowanych metod statystycznych.

Dwie zmienne mogą być powiązane zależnością funkcyjną lub zależnością statystyczną (korelacyjną).

Związek funkcyjny odznacza się tym, że każdej wartości jednej zmiennej niezależnej (będziemy ją oznaczać jako X) odpowiada tylko jedna, jednoznacznie określona wartość zmiennej zależnej (Y). Wiadomo na przykład, że obwód kwadratu jest funkcją jego boku.

Związek statystyczny polega na tym, że określonym wartościom jednej zmiennej odpowiadają ściśle określone średnie wartości drugiej zmiennej. Można zatem obliczyć, jak się zmieni (średnio biorąc) wartość zmiennej zależnej Y w zależności od wartości zmiennej niezależnej X. Oczywiście najpierw na podstawie analizy merytorycznej należy logicznie uzasadnić występowanie związku, a dopiero potem przystąpić do określenia siły i kierunku zależności. Znane są bowiem w literaturze badania zależności (nawet istotnej statystycznie) między liczbą zajętych gniazd bocianich a liczbą urodzeń na danym obszarze czy między liczbą zarejestrowanych odbiorników TV a liczbą chorych umysłowo.

Zwróćmy też uwagę, że liczbowe stwierdzenie występowania zależności nie zawsze oznacza występowanie związku przyczynowo-skutkowego między badanymi zmiennymi. Współwystępowanie dwóch zjawisk może również wynikać z bezpośredniego oddziaływania na nie jeszcze innego trzeciego zjawiska.

Kiedy dwie zmienne są powiązane (skorelowane) możemy na podstawie jednej z nich przewidywać drugą, jednak nie możemy wnioskować co jest przyczyną ich związku.

Ekstrawersja: Emocjonalność w zakresie emocji pozytywnych. Czy towarzyskość powoduje większą satysfakcję z życia? Model korelacyjny nie odpowie na to pytanie. Związek przyczynowo-skutkowy obustronny, dodatkowo sprzężenie zwrotne 43 min nagrania wykładu.

Niska samoocena a neurotyczność. Czasem jest przedstawian jako skądinąd, osoby n. częściej gorzej wypadają na tle prac społecznych, o. silnie reagując są mniej w stanie efektywnie wykonywać zadania, a więc ich brak efektywności może prowadzić do mniejszej samooceny i niskiej skuteczności. Prawo Yerkesa-Dobsona w kontekście osób neurotycznych i niskiej efektywności.

Korelacja nie wskazuje na przyczynowość! Rzeczona korelacja występuje na przykład pomiędzy towarzyskością a satysfakcją z życia. Związek przyczynowo-skutkowy może być też dwustronny – zarówno towarzyskość wpływa na większą satysfakcję z życia, jak i satysfakcja z życia pociąga za sobą większą towarzyskość. Znając jedynie korelację pomiędzy dwiema zmiennymi, nie można prawidłowo określić kierunku zależności. Możliwa jest również jeszcze inna interpretacja przyczynowo-skutkowa. Niewykluczone, że to trzecia zmienna – posiadanie wielu przyjaciół – skłania ludzi do większej towarzyskości i prowadzi do wyższej satysfakcji z życia.

Kiedy związek pomiędzy dwiema zmiennymi można wyjaśnić za pomocą trzeciej zmiennej, mówimy o zależności pozornej.

Moderator to trzecia zmienna która wpływa na związek między dwoma zmiennymi; odpowiada na pytanie kiedy działa dana zależność, kiedy rejestrujemy, którą eksponujemy w danym współczynniku korelacji. Zależność może być pozytywna, ale jak uwzględnimy moderatora, to okazuje się, że dla jednej wartości moderatora ta wartość jest pozytywna, a dla innej negatywna, albo słaba/silna.

Np. stres w postaci nadmiernej stymulacji: nadmierna stymulacja wpływa na napięcie emocjonalne, czyli natężenie emocji negatywnych; przy temperamencie dotyczy to u osób wysoko reaktywnych.

Prawo Yerkesa Dodsona to zależność krzywoliniowa między poziomem motywacji, poziomem napięcia emocjonalnego a efektywnością działania

W analizie korelacji badacz jednakowo traktuje obie zmienne – nie wyróżniamy zmiennej zależnej i niezależnej. Korelacja między X i Y jest taka sama, jak między Y i X. Mówi nam ona, na ile obie zmienne zmieniają się równocześnie w sposób liniowy.

Definicja: korelacja między zmiennymi X i Y jest miarą siły liniowego związku między tymi zmiennymi.

Analizę związku korelacyjnego między badanymi cechami rozpoczynamy zawsze od sporządzenia wykresu.

Wykresy, które reprezentują obrazowo związek pomiędzy zmiennymi, nazywane są wykresami rozrzutu (scatterplot)- „zrób scatterplota” □

Wzrokowa ocena ułatwia określenie siły i rodzaju zależności. Przyjmijmy, że zbiorowość jest badana ze względu na dwie zmienne X i Y, a wartości tych zmiennych w populacji lub próbie n-elementowej są zestawione w postaci dwóch szeregów szczegółowych lub rozdzielczych.

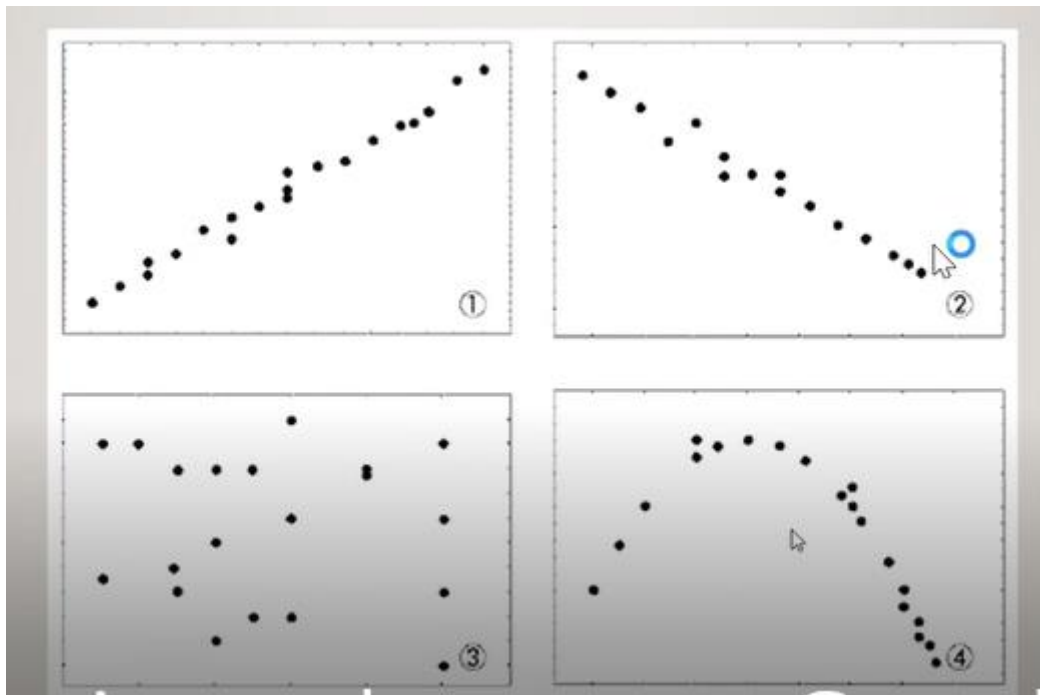
W prostokątnym układzie współrzędnych na osi odciętych zaznaczamy wartości jednej zmiennej, a na osi rzędnych – wartości drugiej zmiennej. Punkty odpowiadające poszczególnym wartościom cech tworzą korelacyjny wykres rozrzutu. Rzadko się zdarza, że zaznaczone punkty leżą dokładnie na linii prostej (pełna korelacja); częściej spotykana konfiguracja składa się z wielu zaznaczonych punktów leżących mniej więcej wzdłuż konkretnej krzywej (najczęściej linii prostej).

Rys. 1 korelacja dodatnia (od lewego dolnego rogu, do prawego górnego rogu)

Rys.2 korelacja ujemna (od lewego górnego rogu, do prawego dolnego rogu)

Rys.3 brak korelacji, współczynnik korelacji $r=0$

Rys. 4 $r=0$, bo jest to współczynnik korelacji liniowej, a nie krzywoliniowej, ale musimy się wtedy zastanowić czy zjawisko które badamy jest związkiem krzywoliniowym.



Jeżeli przyjmiemy, że obie zmienne X i Y są zmiennymi losowymi, to badanie związku między nimi to analiza korelacji. Silnie skorelowane zmienne zachowują się tak, jakby równocześnie się poruszały. Siłę (stopień) korelacji mierzy współczynnik korelacji.

Najbardziej popularny jest współczynnik korelacji liniowej Pearsona, oznaczony symbolem r_{xy} i przyjmujący wartości z przedziału $[-1,1]$. Należy zwrócić uwagę, że współczynnik korelacji Pearsona wyliczamy wówczas, gdy obie zmienne są mierzalne i mają rozkład empiryczny zbliżony do normalnego, a zależność jest prostoliniowa (stąd nazwa).

Korelacje -> parami.

taub Kendalla nieparametryczny, dla zmiennych porządkowych

Korelacje			
		AFE	Cynizm
AFE	Korelacja Pearsona	1	-.515**
	Istotność (dwustronna)		.000
	Suma kwadratów i iloczyn wektorowe	3132.657	-2774.965
	Kowariancja	22.061	-19.542
	N	143	143
Cynizm	Korelacja Pearsona	-.515**	1
	Istotność (dwustronna)	.000	
	Suma kwadratów i iloczyn wektorowe	-2774.965	9265.119
	Kowariancja	-19.542	65.247
	N	143	143

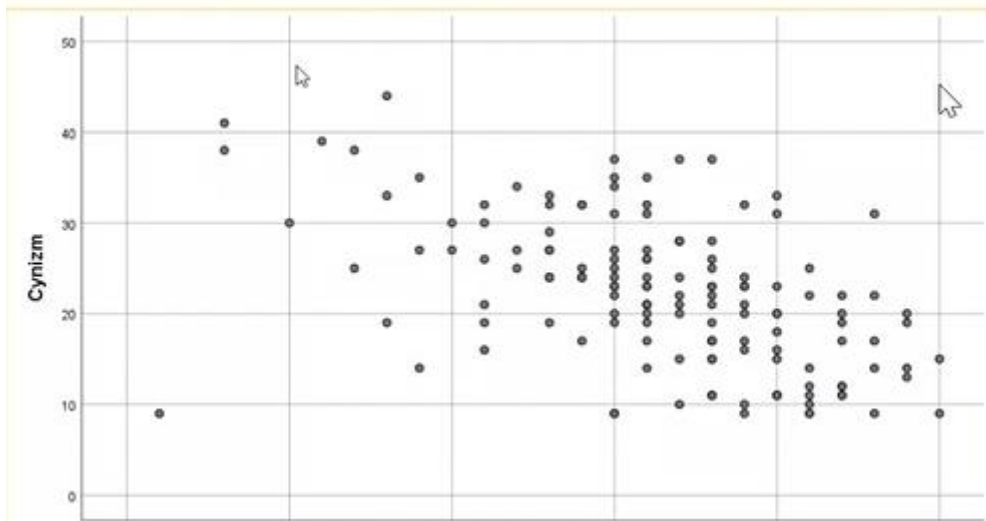
** Korelacja istotna na poziomie 0.01 (dwustronnie).

Afektywne przywiązanie do organizacji a acynizm organizacyjny (negatywny stosunek wobec organizacji). Korelacja ujemna -.51**

Im wyższe przywiązanie afektywne tym mniejszy cynizm. Korelacja jest istotna gdy odstrzegamy wiele zer po przecinku. 1h 13 min nagrania

Wykresy -> tradycyjne -> rozrzutu/punktowe

Przewidywanie natężenia cynizmu organizacyjnego na podstawie afektywnego przywiązania. Oś X to AFE oś y cynizm.



1:15

Przy interpretacji współczynnika korelacji liniowej Pearsona należy więc pamiętać, że wartość współczynnika bliska zero nie zawsze oznacza brak zależności, a jedynie brak zależności liniowej

ZMIENNA NIEZALEŻNA

Zmienna niezależna (zmienna wyjaśniająca) – zmienna, którą kontrolujemy i której związek/wpływ ze zmienną zależną chcemy wykazać.

W przypadku badań [korelacyjnych](#) podział na zmienną zależną i niezależną jest umowny, wiele źródeł podaje, że takie ich określanie jest błędem. Niemniej, czasami w analizie korelacji traktuje się te zmienne tak, jak gdyby wprowadzane były do modelu [regresji](#), uważając predyktory za zmienne niezależne, a zmienną wyjaśnianą, jako zależną.

Zmienna wyjaśniana pojawia się na osi Y, a wyjaśniającą na osi X

r = współczynnik korelacji w ramach próby ρ = współczynnik korelacji w populacji

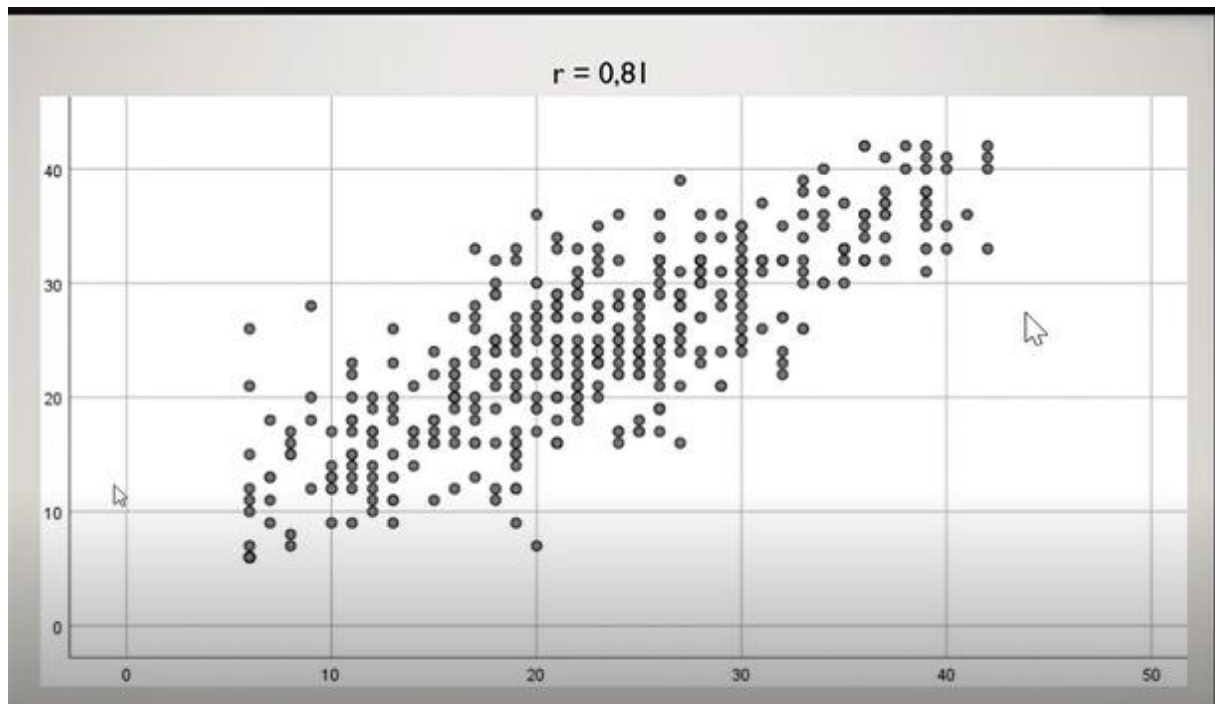
1. Gdy $\rho=0$ nie ma korelacji, czyli nie ma liniowego związku między dwiema zmiennymi losowymi. Gdy $\rho=1$, zachodzi ścisły związek między dwoma zmiennymi. Znaczy to, że gdy jedna z tych zmiennych przyjmuje większe wartości, to i druga przyjmuje większe wartości, a gdy wartości jednej

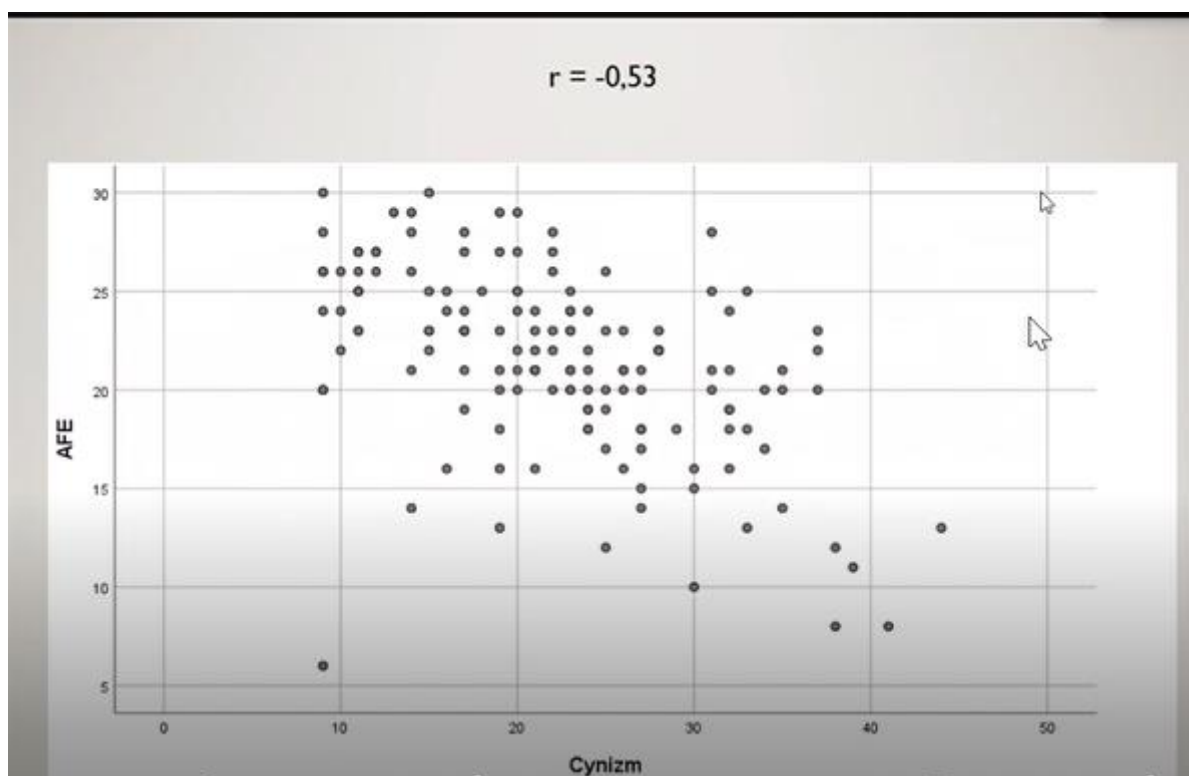
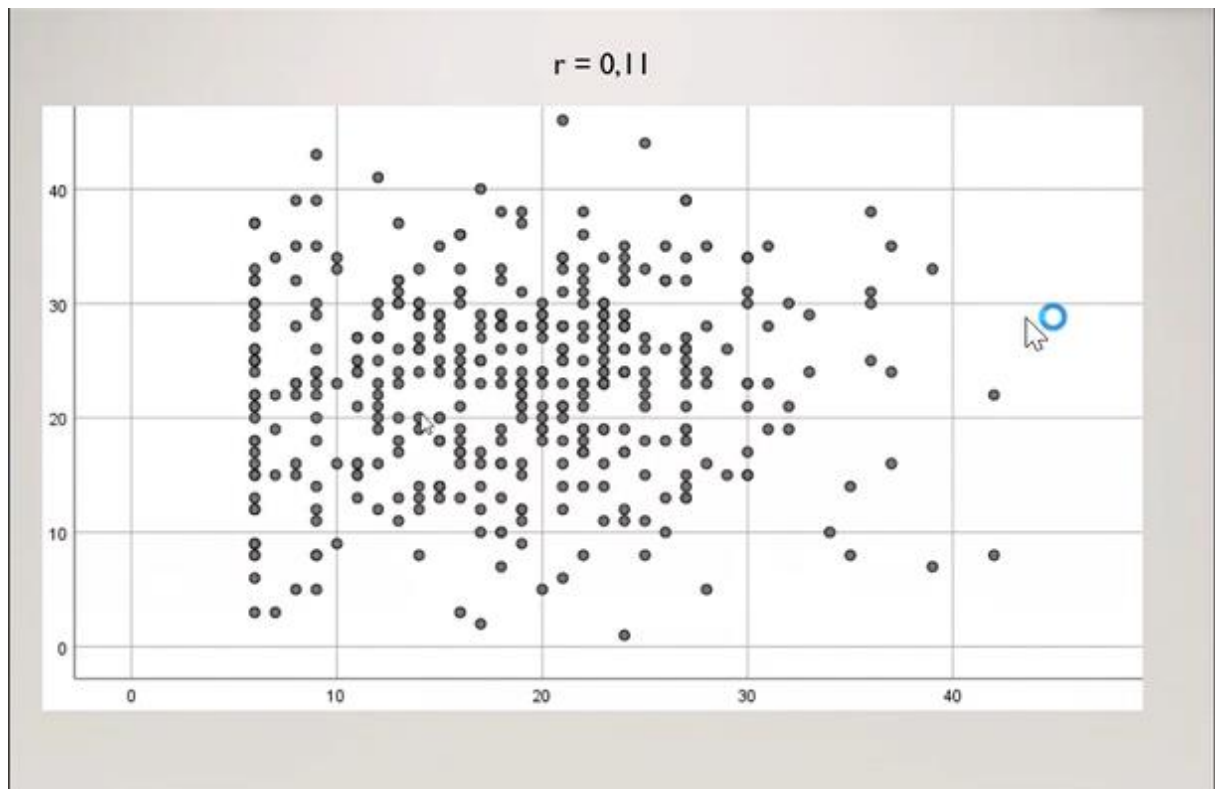
zmiennej spadają, to i wartości drugiej zmiennej spadają.

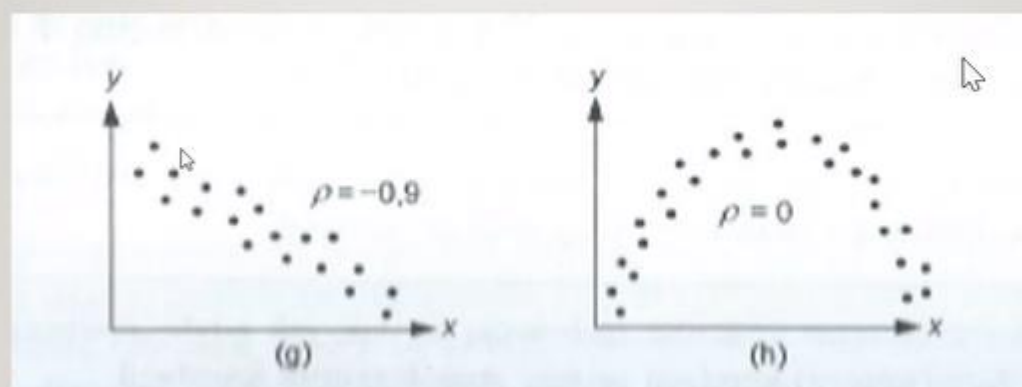
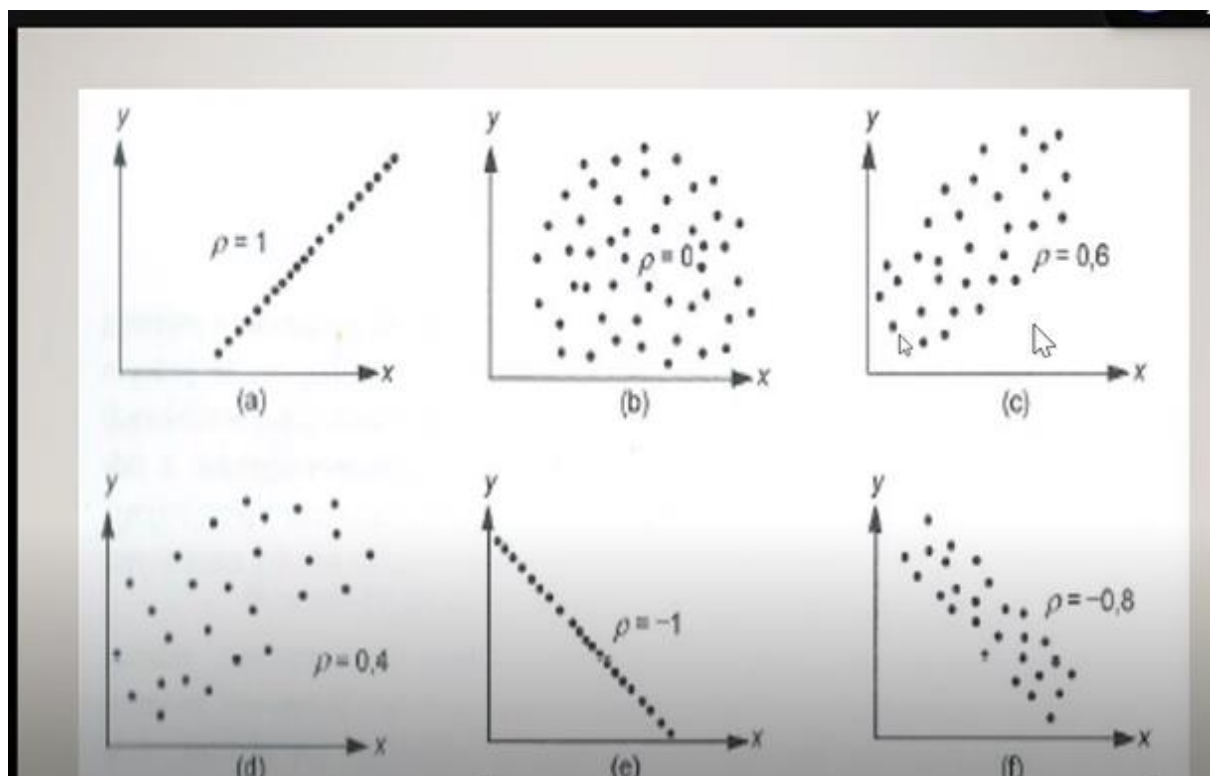
Gdy $p = -1$ zachodzi ścisły związek między dwiema zmiennymi. Gdy jedna z tych zmiennych przyjmuje większe wartości, to druga przyjmuje mniejsze wartości, a gdy wartości jednej zmiennej spadają, to wartości drugiej zmiennej rosną.

Jeśli bezwzględna wartość p mieści się wewnątrz przedziału $[0,1]$, to między ona siłą liniowego związku między dwiema zmiennymi.

Np. Wartość współczynnika korelacji 0,90 oznacza silny dodatni liniowy związek między zmiennymi. Wartość -0,70 współczynnika korelacji oznacza nieco słabszy i ujemny związek liniowy. Wartość współczynnika korelacji 0,30 oznacza słaby związek liniowy między zmiennymi.

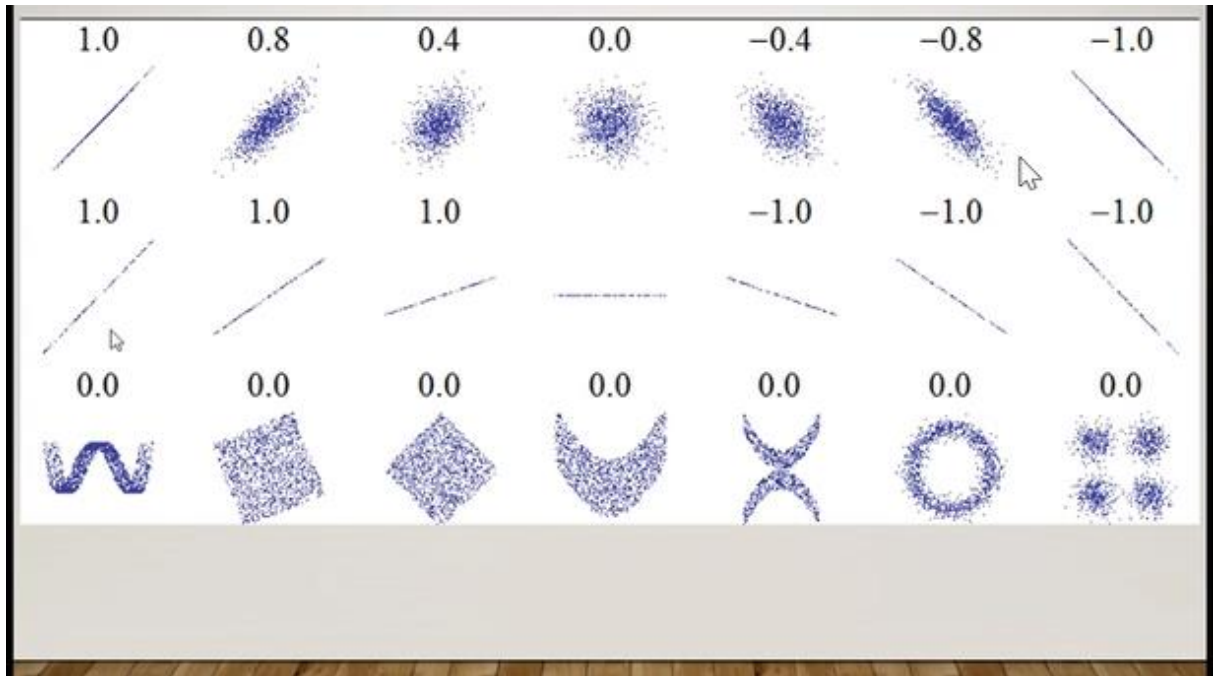






Czy determinacja i wspólna wariancja są używane zamiennie? To jest to samo?

W nawiązaniu do teorii matematycznej powinniśmy mówić o współczynniku determinacji i o tym na ile wariancja zmiennej x jest wyjaśniona przez wariancję zmiennej y . Jednakże w wielu publikacjach można spotkać się z terminem „wspólna wariancja”



W analizie statystycznej zwykle przyjmuje się następującą skalę:

$R_{xy} = 0$ zmienne nie są skorelowane

$0 < r_{xy} < 0,1$ korelacja nikła

$0,1 = < r_{xy} < 0,3$ korelacja słaba

$0,3 = < r_{xy} < 0,5$ korelacja przeciętna

$0,5 = < r_{xy} < 0,7$ korelacja wysoka

$0,7 = < r_{xy} < 0,9$ korelacja bardzo wysoka

$0,9 = < r_{xy} < 1$ korelacja prawie pełna

W jaki sposób dochodzimy do pojęcia korelacji? Rozpatrzmy parę zmiennych losowych X i Y zakładając, że zarówno X jak i Y mają rozkłady normalne o średnich μ_x i μ_y oraz o standardowych odchyleniach σ_x i σ_y . Zdefiniujemy kowariancję zmiennych X i Y .

Kowariancja dwóch zmiennych X i Y :

$\text{cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$, gdzie μ_x i μ_y są średnimi zmiennych X i Y w populacji,

Kowariancja jest więc oczekiwaną wartością iloczynu odchylenia wartości zmiennych X i Y od ich średnich. Kowariancja przyjmuje wartość dodatnią, gdy obie zmienne „poruszają się razem w tym samym kierunku”, a wartość ujemną, gdy poruszają się w przeciwnych kierunkach”.

Nie można jej interpretować jako wskaźnika stopnia liniowego związku między zmiennymi, ponieważ jej wielkość zależy od wielkości standardowych odchyłeń zmiennych X i Y . Jeżeli jednak podzielimy kowariancję przez odchylenia standardowe obu zmiennych, to otrzymamy miarę, która przyjmuje wartości z przedziału od -1 do 1 i informuje nas o sile liniowego związku między zmiennymi. Tą miarą jest właśnie współczynnik korelacji (w populacji).

Kroki obliczania kowariancji:

1. Obliczamy średnie dla obu zmiennych.
2. Odejmujemy wynik osoby w danej zmiennej od średniej dla tej zmiennej. Obliczamy więc odległości wyników w danej zmiennej od jej średniej.
3. Dla każdej osoby mnożymy obie odległości wyników zmiennych od ich średnich.
4. Dodajemy do siebie iloczyny odległości wyników zmiennych od ich średnich. Dodajemy do siebie iloczyny odległości wyników zmiennych od ich średnich. Dodajemy do siebie iloczyny odległości – to jest licznik kowariancji.
5. By uzyskać wartość kowariancji dzielimy obliczoną w kroku 4 sumę przez liczbę obserwacji pomniejszoną o 1

Zbliżenie - kowariancja

Tabela 1.1. Kolejne kroki obliczania wielkości kowariancji dla zmiennych X oraz Y

Wartości zmiennej X	Wartości zmiennej Y	Odległość od średniej dla X	Odległość od średniej dla Y	Iloczyn odległości
1	5	-2	2	-4
2	4	-1	1	-1
3	3	0	0	0
4	2	1	-1	-1
5	1	2	-2	-4
średnia = 3	średnia = 3			suma: -10

Neo-FFI podręcznik:

CECHY OSOBOWOŚCI MIERZONE NEO-FFI A CECHY TEMPERAMENTU
MIERZONE FCZ-KT

Skale FCZ-KT	NEU	EKS	OTW	UGD	SUM
Żwawość	-0,39**	0,28**	0,11**	0,11**	0,32**
Perseweratywność	0,49**	-0,06**	0,09**	-0,03	-0,06**
Wrażliwość sensoryczna	-0,08**	0,08**	0,30**	0,09**	0,12**
Reaktywność emocjonalna	0,64**	-0,29**	-0,13**	-0,02	-0,15**
Wytrzymałość	-0,44**	0,24**	0,10**	0,07**	0,20**
Aktywność	-0,20**	0,63**	0,25**	-0,19**	-0,02

FCZ-KT - Formalna Charakterystyka Zachowania - Kwestionariusz Temperamentu.

** - $p < 0,01$, * - $p < 0,05$.

CECHY OSOBOWOŚCI MIERZONE NEO-FFI A CECHY TEMP
MIERZONE EAS-TS

Skale EAS-TS	NEU	EKS	OTW	UGD
Towarzyskość	-0,20**	0,55**	0,00	0,11**
Aktywność	-0,10**	0,51**	0,14**	-0,05*
Strach	0,59**	-0,25**	-0,07**	-0,02
Niezadowolnienie	0,66**	-0,26**	-0,08**	-0,19**
Gniew	0,42**	-0,01	-0,04	-0,33**

EAS-TS(D) - Kwestionariusz Temperamentu EAS-TS wersja dla o
(*Emotionality Activity Sociability - Temperament Survey*).

** - $p < 0,01$, * - $p < 0,05$.

CECHY OSOBOWOŚCI MIERZONE NEO-FFI A CECHY TEMP MIERZONE EAS-TS

Skale EAS-TS	NEU	EKS	OTW	UGD
Towarzyskość	-0,20**	0,55**	0,00	0,11**
Aktywność	-0,10**	0,51**	0,14**	-0,05*
Strach	0,59**	-0,25**	-0,07**	-0,02
Niezadowolenie	0,66**	-0,26**	-0,08**	-0,19**
Gniew	0,42**	-0,01	-0,04	-0,33**

EAS-TS(D) - Kwestionariusz Temperamentu EAS-TS wersja dla o
(*Emotionality Activity Sociability - Temperament Survey*).

** - $p < 0,01$, * - $p < 0,05$.

Korki obliczania współczynnika korelacji r Pearsona:

1. Obliczamy średnie i odchylenia standardowe dla obu zmiennych.
2. Standaryzujemy wyniki każdej zmiennej, odejmując of każdego wyniku średnią i dzieląc tę różnicę przez odchylenie standardowe.
3. Dla każdej osoby mnożymy wystardaryzowane wyniki dla obu zmiennych.
4. Dodajemy do siebie iloczyny wystardaryzowanych wyników – to jest licznik współczynnika korelacji r Pearsona.
5. Aby uzyskać wartość korelacji, dzielimy obliczoną w kroku 4. Sumę przez liczbę obserwacji pomniejszoną o 1.

Tabela 1.2. Kolejne kroki obliczania wielkości korelacji dla zmiennych X oraz Y

Wartości zmiennej X	Wartości zmiennej Y	Wystandaryzowana odległość od średniej dla X ($(X_i - M)/SD$)	Wystandaryzowana odległość od średniej dla X ($(X_i - M)/SD$)	Iloczyn odległości
1	5	-1,26	1,26	-1,6
2	4	-0,63	0,63	-0,4
3	3	0,00	0,00	0,0
4	2	0,63	-0,63	-0,4
5	1	1,26	-1,26	-1,6
średnia = 3 SD = 1,6	średnia = 3 SD = 1,6			suma: -4

Tabela do kroku 4.

W 4 kolumnie jest błąd, powinien być „dla Y ”

$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$

WSPÓŁCZYNNIK KORELACJI RANGOWEJ SPEARMANA

$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$

gdzie: D – różnica między parą rang,
 n – liczba par rang.

Często stosowanym estymatorem jest

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

gdzie (tak jak wcześniej)

$$d_i = R_{x_i} - R_{y_i}$$

to różnica między rangami zmiennych X i Y dla obserwacji i

Nieparametryczny odpowiednik współczynnika korelacji Pearsona

Osoba badana	① Porządek zwracania testów X	② Wynik testu Y	③ Rangi dla X R_x	④ Rangi dla Y R_y	⑤ $D = R_x - R_y$	⑥ D^2
A	1	28	1	6,5	-5,5	30,25
B	2	21	2	2	0,0	0,00
C	3	22	3	3,5	-0,5	0,25
D	4	22	4	3,5	0,5	0,25
E	5	32	5	10	-5,0	25,00
F	6	36	6	13	-7,0	49,00
G	7	33	7	11	-4,0	16,00
H	8	39	8	15	-7,0	49,00
I	9	25	9	5	4,0	16,00
J	10	30	10	8	2,0	4,00
K	11	20	11	1	10,0	100,00
L	12	28	12	6,5	5,5	30,25
M	13	31	13	9	4,0	16,00
N	14	38	14	14	0,0	0,00
O	15	34	15	12	3,0	9,00
$n = 15$						$\sum D^2 = 345,00$

Obliczenia:

$$\textcircled{7} r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)} = 1 - \frac{6(345)}{15(15^2 - 1)} = 0,38$$

Tak jak wartość innych parametrów populacji współczynnik korelacji (w populacji) nie jest znany i musimy go oszacować na podstawie znajomości losowej próby par wyników obserwacji zmiennych X i Y. Tak wyliczony z próby współczynnik r_{xy} jest estymatorem współczynnika korelacji ρ w populacji generalnej, a jego wartość liczbową stanowi ocenę punktową siły powiązania w całej populacji. Stąd konieczność testowania istotności współczynnika korelacji wyliczonego w oparciu o próbę losową. Najpowszechniej stosowany test polega na sprawdzeniu, czy zmienne X i Y są w ogóle skorelowane.

Współczynnik r jest próbą oszacowania współczynnika korelacji w populacji generalnej.

To jest ten test

WARTOŚĆ t POTRZEBNA DO TESTOWANIA HIPOTEZY ZEROWEJ O POSTACI $\rho = 0$

$$t = \frac{r}{\sqrt{(1-r^2)/(n-2)}}$$

gdzie: r = współczynnik z próby,
 n = liczba par wyników w próbie

$H_0: \rho = 0$

$H_1: \rho \text{ różne od } 0$

Weryfikacja tej hipotezy zerowej pomoże nam w ocenie, czy istniejąca zależność między X i Y w próbie jest tylko przypadkowa, czy też jest prawidłowością w populacji. Dlatego test istotności współczynnika korelacji Pearsona spotykamy we wszystkich pakietach statystycznych.

Wyliczamy zwykły test t, który nas informuje o tym czy daną hipotezę zerową utrzymać czy ją odrzucić. Jeśli ją odrzucamy, co widzieliśmy dzisiaj, gdy korelacja jest ogwiazdkowana, to wówczas odrzucamy hipotezę zerową i stwierdzamy że współczynnik korelacji jest różny, istotnie różny od 0. Czyli między dwiema zmiennymi zachodzi zależność liniowa.

Wykład 18.04 Grabowski – przykłady modelu korelacyjnego, badania sondażowe, kwestionariuszowe

Alfa cronbacha jest wskaźnikiem rzetelności danego testu; spójność wewnętrzna danego kwestionariusza, skali. Obliczmy zaznaczając pozycje jednej i tej samej skali. Im bardziej są one skorelowane ze sobą tym większą wartość przyjmuje alfa cronbacha.

Metoda połówkowa – dzielimy skalę na dwie części i sprawdzamy jak są ze sobą skorelowane

Pomiar w psychologii

Najczęściej dobierane wskaźniki, których obecność wskazuje na natężenie danej cechy. W ten sposób cecha latentna będzie w procesie operacjonalizacji reprezentowana przez określone, możliwe do zaobserwowania zjawiska.

Dokonanie pomiaru musi zostać przedstawione na jakiejś skali pomiarowej.

Metoda korelacyjna

Metoda, w której mierzy się systematycznie dwie zmienne lub ich większą liczbę i oszacowuje się relację pomiędzy nimi (tj. w jakim stopniu można przewidzieć wartości jednej zmiennej na podstawie wartości drugiej zmiennej).

Badania za pomocą skal (skala Likerta 1-5)

Skale [arametryczne badania parametry populacji.

Korelacja dodatnia – taka relacja pomiędzy dwiema zmiennymi, w której wzrostowi wartości jednej zmiennej towarzyszy wzrost wartości drugiej zmiennej.

Korelacja ujemna – taka relacja pomiędzy dwiema zmiennymi, w której wzrostowi wartości jednej zmiennej towarzyszy zmniejszanie się (spadek) wartości drugiej zmiennej.

Statystyczny.pl

Stawianie i weryfikacja hipotez to wnioskowanie statystyczne.

Podstawowa hipoteza (hipoteza zerowa) w odniesieniu do współczynnika korelacji jest hipoteza, która mówi o tym, że współczynnik korelacji przyjmuje wartość zero – wówczas korelacja nie istnieje (nie istnieje między nimi związek liniowy). Weryfikacja hipotezy zerowej pomaga nam w ocenie czy istniejąca zależność między zmienną x a y jest tylko przypadkowa, czy też jest prawidłowością w populacji. J

Hipoteza alternatywna głosi o tym, że r (ro) jest różna od zera, czyli istnieje związek liniowy między zmiennymi.

Jeżeli odrzucamy hipotezę zerową i przyjmujemy hipotezę alternatywną, to znaczy, że jest to pewna prawidłowość w populacji. Dzieje się tak zazwyczaj, gdy dysponujemy próbą tzn. reprezentatywną, która odnosi się do struktury cech w populacji.

The diagram illustrates the process of testing the null hypothesis for a Pearson correlation coefficient. It features a central formula for the test statistic t , which is the sample correlation coefficient r divided by the square root of $(1-r^2)/(n-2)$. A red arrow points from the text 'To jest ten test' to the formula. Above the formula, it states 'WARTOŚĆ : POTRZEBNA DO TESTOWANIA HIPOTEZY ZEROWEJ O POSTACI $\rho = 0$ '. Below the formula, it defines r as the sample correlation coefficient and n as the number of pairs of results in the sample. To the left, the null hypothesis $H_0: \rho = 0$ and the alternative hypothesis $H_1: \rho \text{ różne od } 0$ are listed. At the bottom, a paragraph explains that this test helps determine if a linear relationship between X and Y in the sample is merely coincidental or a true property of the population, and that this test is commonly found in statistical software packages.

To jest ten test

WARTOŚĆ : POTRZEBNA DO TESTOWANIA HIPOTEZY ZEROWEJ O POSTACI $\rho = 0$

$$t = \frac{r}{\sqrt{(1-r^2)/(n-2)}}$$

gdzie: r = współczynnik z próby,
 n = liczba par wyników w próbie

$H_0: \rho = 0$

$H_1: \rho \text{ różne od } 0$

Weryfikacja tej hipotezy zerowej pomoże nam w ocenie, czy istniejąca zależność między X i Y w próbie jest tylko przypadkowa, czy też jest prawidłowością w populacji. Dlatego test istotności współczynnika korelacji Pearsona spotykamy we wszystkich pakietach statystycznych.

n to liczba obserwacji, ten wzór można przenieść do excela i obliczyć; ρ to współczynnik korelacji w populacji

W pracy magisterskiej często stosujemy dobór celowy. Dobór losowy jest trudny technicznie

n:	423
r:	0,799278
r ² :	0,638845321
1-r ² :	0,361154679
n-2:	421
	0,00085785
	0,02928907
t:	27,28929301
p=	0,000000000000000000000000
0,05	1,965614792
0,01	2,587557573
0,001	3,313785889
0,0001	3,890622949

test t służący do weryfikacji hipotezy zerowej.

Wzór z grafiki powyżej rozłożony na części pierwsze

n liczba par (osob badanych)

r = współczynnik korelacji (wartość)

r^2 = współczynnik determinacji

$1-r^2$ =współ indeterminacji

(odpowiada na pytanie „jaka część

wariancji nie jest wyjaśniania przez

zmienną x w przypadku zmiennej y)

$n-2$ = ?

t = wartość testu t

p = weryfikowana jest jego istotność

(poniżej różne poziomy istotności –

0,05;0,01 itd.)

Korelacje [new1.sta]										
Oznaczone wsp. korelacji są istotne z $p < ,05000$										
Srednia	Odch. St d	$r(X,Y)$	r_s	t	p	N	Stała zał: Y	Nachyle zał: Y	Stała zał: X	Nachyle zał: X
7,46	5,88									
95,42	15,04	,731	,534	3,550	,005	13	81,48	1,869	-19,8	,286
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]

5. wartość statystyki t badającej istotność współczynnika korelacji

6. Poziom istotności

7. Liczebność grupy

8. Wyraz wolny regresji liniowej Y względem X (predyktorem jest zmienna Y)

9. Współczynnik regresji liniowej zmiennej Y względem zmiennej X (predyktorem jest zmienna Y)

10. Wyraz wolny regresji liniowej względem Y (predyktorem jest zmienna X)

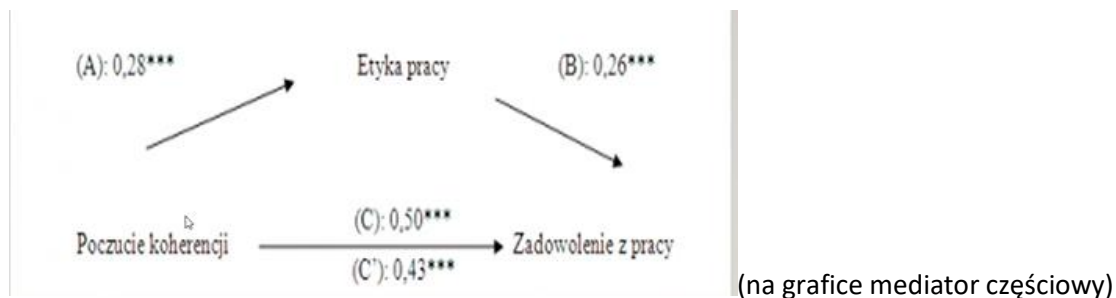
11. Współczynnik regresji liniowej zmiennej x względem zmiennej Y . (predyktorem jest zmienna X)

Do wnioskowania o zależnościach przyczynowo-skutkowych trzeba czegoś więcej niż tylko danych korelacyjnych. Może to być wyszukanie trzeciej zmiennej.

Mediatory i moderatory (trzecia zmienna Z),

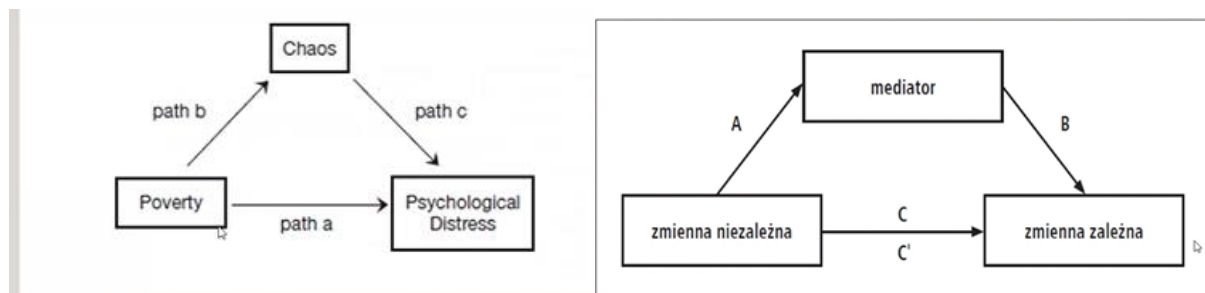
mediatory związane są z perspektywą analiz ścieżkowych. Analizy ścieżkowe to metody statystyczne

pomocne w rozumieniu potencjalnych przyczyn związków korelacyjnych. Pomaga nam to zrozumieć korelację pozorną.



Jeżeli współczynnik ścieżki C spada do zera albo staje się bardzo niski wówczas możemy powiedzieć, że zmienna Z (etyka pracy) jest mediatorem całkowitym.

Mediatory są zmiennymi pomocnymi w wyjaśnianiu związków pomiędzy dwiema zmiennymi. W analizach mediacji nazwy zależna i niezależna stosowane są tylko umownie.



Ubóstwo jako zmienna niezależna, chaotyczny styl życia mediator, dystres psychologiczny zmienna zależna (bo zależy od zmiennej niezależnej). Zbliżenie jak wykonać analizę ścieżek (rysunek 2).

Jak wykonać analizę ścieżek? Kroki analizy mediacji

1. Sprawdzenie czy relacja między zmienną niezależną a zależną jest istotna (ZN-ZZ) – podstawowa regresja
2. Sprawdzenie czy relacja między zmienną niezależną a mediatorem jest istotna (ZN-M)
3. Sprawdzenie czy relacja między zmienną niezależną po wprowadzeniu mediatora do modelu regresji jest słabsza; wykonanie dodatkowego testu sprawdzającego istotność iloczynu współczynników regresji dla pośredniej relacji między zmienną niezależną a zależną z uwzględnieniem mediatora. – regresja uwzględniająca zmienną niezależną i mediatora, zmienną wyjaśnianą jest zmienna zależna (nazwy zmiennych w dalszym ciągu umowne).

Mediatory to często zmienne dychotomiczne

Kwestionariusze

Większość badań ankietowych do pomiaru zmiennych wykorzystuje kwestionariusze. Skale samoopisu stosuje się do mierzenia opinii i różnic indywidualnych między ludźmi. Skale te mierzą myśli i uczucia dotyczące różnych aspektów rzeczywistości. Przez pryzmat tych myśli, odczuć mierzone są takie cechy jak temperament oraz postawy.

Do pomiaru różnic między respondentami służą dwie metody: skale Likerta i analiza czynnikowa.

Kwestionariusze powinny być trafne i rzetelne, co wymaga umiejętności ich konstruowania.

Przykład: NEO FFI

Błąd pomiaru

Na błąd pomiaru wpływa wielkość próby i poziom ufności, który wybiera badacz przy szacowaniu wyników populacji. Choć błąd pomiaru wskazuje jak dalece wyniki sondażu mogą opisywać populację, możliwość interpretacji wyniku zależy od jakości metod wykorzystywanych w badaniu.

Możliwe jest oszacowanie błędu pomiaru, czyli różnicy pomiędzy wynikiem próby i prawdziwymi wartościami populacji.

Błąd pomiaru podaje raczej przedział, w którym znajduje się prawdziwa wartość populacji (np. pomiędzy 60% a 70% populacji preferuje obecną panią burmistrz), aniżeli szacuje dokładną jej wartość (np. 65% populacji preferuje obecną burmistrz).

Jeśli chodzi o błąd standardowy to wielkość próby wyznacza błąd pomiaru. Im większa próba, tym mniejszy błąd pomiaru. Większa próba lepiej reprezentuje różne właściwości populacji.

Zatem, aby zmniejszyć błąd pomiaru (i bardziej precyzyjnie opisać populację) należy zwiększyć wielkość próby.

Znajomość błędu pomiaru nie wystarcza jednak do uzyskania pewności, że wyniki sondażu w ogóle można interpretować.

Możliwość taka zależy również od tego typu czynników jak:

1. Sposób doboru próby
2. Liczba osób z wybranej próby, które odpowiedziały na pytania sondażu.
3. Sposób formułowania pytań oraz
4. Przeszkolenie ankietatorów zbierających odpowiedzi

Rzetelność

Rzetelność odnosi się do wewnętrznej spójności pomiaru, która często jest oceniana za pomocą ponownego pomiaru tej samej próby. Rzetelność rośnie, gdy wzbogacamy dane narzędzie o podobne pozycje, badamy zróżnicowane próby osób i stosujemy ujednolicone procedury.

Kolejnym dobrym sposobem rzetelności jest ponowne jej użycie. Mierzmy skalą w jednym momencie czasowym i po kilku tygodniach używamy narzędzia na tej samej próbie i badamy korelację.

Stosowanie jednego pytania do jednej zmiennej pozbawia nas rzetelności pomiaru. Min. 3 itemy!

Podobne pozycje zwiększają spójność wewnętrzną danej skali. Im bardziej podobne do siebie pozycje, tym większy wskaźnik alfa Cronbacha.

Wskaźniki rzetelności (spójność wewnętrzną).

Rzetelność kwestionariusza WPEP

Tabela 4

Wskaźniki rzetelności (α Cronbacha) siedmiu skal WPEP dla prób A i B oraz studentów (St)

Próba	CP	PS	NCW	PC	ME	OG	NMC
St	$\alpha = 0,85$	$\alpha = 0,85$	$\alpha = 0,89$	$\alpha = 0,82$	$\alpha = 0,76$	$\alpha = 0,75$	$\alpha = 0,73$
A	$\alpha = 0,86$	$\alpha = 0,84$	$\alpha = 0,88$	$\alpha = 0,82$	$\alpha = 0,75$	$\alpha = 0,77$	$\alpha = 0,77$
B	$\alpha = 0,83$	$\alpha = 0,81$	$\alpha = 0,88$	$\alpha = 0,79$	$\alpha = 0,67$	$\alpha = 0,73$	$\alpha = 0,75$

Źródło: opracowanie własne.

Współczynniki korelacji pomiaru po pięciu tygodniach z pomiarem wyjściowym wyniosły dla poszczególnych skal CP $r = 0,79$, PS $r = 0,85$, NCW $r = 0,81$, PC $r = 0,81$, ME $r = 0,83$, OG $r = 0,85$, NMC $r = 0,78$ (wszystkie współczynniki korelacji istotne na poziomie $p = 0,001$). Wskazuje to na wysoką stabilność pomiaru skalami WPEP.

https://www.researchgate.net/profile/Damian-Grabowski/publication/291356277_Wielowymiarowy_Profil_Etyki_Pracy_WPEP_Charakterystyka_psychometryczna_polskie

Grabowski/publication/291356277_Wielowymiarowy_Profil_Etyki_Pracy_WPEP_Charakterystyka_psychometryczna_polskie

wartości powyżej 0,6

alfa cronbacha to zadowalająca (wystarczająca) wartość (rzetelność), w pełni satysfakcjonująca rzetelność to 0,7, wysoka rzetelność poziom 0,8

Trafność odnosi się do prawdziwości pomiaru : stwierdza czy mierzy on to, do mierzenia czego został stworzony.

Trafność teoretyczna oznacza stopień w jakim pozycje testowe odnoszą się do pojęć, które mają mierzyć. Trafność teoretyczna wyznacza trafność zbieżna i trafność różnicowa. Zajmując się trafnością zbieżną, zakładamy, że wyniki kwestionariusza badającego daną cechę będą znacznie korelować z wynikami narzędzia badającego tę samą cechę lub podobną.

Trafność różnicowa – testy badające daną cechę będą korelować niżej lub wcale z wynikami narzędzia badającego inną cechę.

Trafność zbieżna. Wyniki kwestionariusza badającego daną cechę będą znacznie korelować z wynikami narzędzia badającego tę samą cechę lub podobną.

Tabela 5

Wartości współczynników korelacji między religijnością RCI a wymiarami duchowości w kwestionariuszu Samoopisu (N = 343)

	M	SD	PR	WE	H	R-intra	R-inter	RCI
PR	18,12	7,02						
WE	25,94	4,83	0,54***					
H	23,15	4,91	0,55***	0,53***				
R-intra	16,04	6,76	0,89***	0,50***	0,48***			
R-inter	9,94	4,25	0,77***	0,43***	0,42***	0,84***		
RCI	25,98	10,56	0,88***	0,49***	0,47***	0,98***	0,94***	

Uwaga. *** $p < 0,001$.

1. **Postawy religijne (PR)** odnoszą się do emocji, myśli, czyli przeżyć religijnych, ich znaczenia w codziennym życiu i ich wpływu na wybory moralne i postępowanie, wyrażany stosunek do Boga
2. **Wrażliwość etyczna (WE)** oznacza, że w systemie poznawczym osoby wysokie miejsce zajmują wartości moralne (etyczne).
3. **Harmonia (H)** oznacza skłonność osoby do poszukiwania harmonii ze światem, ładu wewnętrznego, spójności różnych form własnej aktywności.

Trafność różnicowa. Testy badające daną cechę będą korelować niżej lub wcale z wynikami narzędzia badającego inną cechę.

Tutaj oczekiwano, że wymiary etyki pracy będą korelowały bardzo słabo z N, E, O, A (trafność różnicowa) a mocniej z C (sumienność) (trafność zbieżna)

Tabela 7

Wartości współczynników korelacji między wskaźnikiem globalnym WPEP oraz wymiarami etyki pracy a pięcioma czynnikami osobowości (NEO-FFI)

N = 515	Neurotyczność	Ekstrawersja	Otwartość	Ugodowość	Sumienność
CP	-0,10*	0,11*	-0,02	-0,04	0,20***
PS	-0,13**	-0,03	0,07	-0,23***	-0,02
NCW	-0,04	-0,06	-0,08	0,04	0,14**
PC	-0,01	0,10*	0,03	0,03	0,30***
ME	-0,11*	0,08	0,08	0,20***	0,19***
OG	-0,07	0,04	-0,07	0,11*	0,13**
NMC	-0,04	0,09*	-0,01	-0,02	0,35***
WPEP	-0,11*	0,08	-0,00	0,02	0,31***
Średnia korelacja	-0,07	0,05	-0,00	0,01	0,20

*** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$.

Źródło: opracowanie własne.

https://www.researchgate.net/profile/Damian-Grabowski/publication/291356277_Wielowymiarowy_Profil_Etyki_Pracy_WPEP_Charakterystyka_psychometryczna_polskiej_wersji_kwestionariusza_Multidimensional_Work_Ethic_Profile/links/56c8ea3d08ae5488f0d7051a/Wielowymiarowy-Profil-Etyki-Pracy-WPEP-Charakterystyka-psychometryczna-polskiej-wersji-kwestionariusza-Multidimensional-Work-Ethic-Profile.pdf

Przewidywanie

przewidywanie aktuarialne to przewidywanie średnich zachowań ludzi na podstawie danych korelacyjnych.

Wykorzystywanie różnorodnych źródeł informacji podnosi trafność przewidywań.

Statystyczne przewidywania na podstawie korelacji przewyższają przewidywania oparte o diagnozę kliniczną.

Dokonywanie przewidywań na podstawie danych korelacyjnych nazywamy przewidywaniem aktuarialnym. Podejmowanie trafnych decyzji uniemożliwiają rzetelne związki między zmiennymi, czyli silne korelacje, na podstawie których można dokładnie przewidywać średnie zachowania ludzi.

Podstawą przewidywań jest analiza regresji.

Badania sondażowe

Reprezentatywność próby. Próba jest reprezentatywna w takim stopniu, w jakim ujawnia podobny rozkład cech jak populacja.

Tendencyjność

Z tendencyjną selekcją mamy do czynienia wtedy, gdy procedury doboru próby powodują nadreprezentację jakiejś warstwy populacji lub odwrotnie – wyłączenia lub niedoreprezentowania jakiejś znaczącej warstwy populacji.

Dobór próby

Istnieją dwie podstawowe metody doboru próby:

1. Nieprobabilistyczny
2. Probabilistyczny

Stosując nieprobabilistycznego, przypadkowego lub celowego doboru próby, nie mamy gwarancji, że każdy element populacji ma jednakową szansę na znalezienie się w próbie.

Ad.1 Nieprobabilistyczny (przypadkowy)

Przypadkowy dobór próby polega na wybraniu respondentów, którzy mogą i chcą brać udział w badaniu. W gazetach często pojawiają się wypowiedzi „osób z ulicy”. Ich opinie bywają interesujące, ale nie mogą być reprezentatywne dla szerszej społeczności ze względu na to, że dobrano ich w sposób przypadkowy. Są to osoby, które np. same zgłaszają się na badania.

Crossen omawia program nadawany w telewizji w czasie największej oglądalności, w którym przeprowadzono sondaż telefoniczny wśród widzów na temat siedziby Narodów Zjednoczonych. Pytanie było następujące: czy siedziba Narodów Zjednoczonych powinna pozostać w USA? Większość (67%) ze 186 tys. Widzów było za przeniesieniem siedziby Narodów Zjednoczonych poza USA. Natomiast większość (72%) respondentów sondażu przeprowadzonego na reprezentatywnej próbie chciało pozostawienia siedziby Narodów Zjednoczonych na terenie USA.

Drugą formą doboru nielosowego stanowi celowy dobór próby, który polega na doborze elementów charakteryzujących się jakąś szczególną właściwością., np. pacjenci depresyjni. Zazwyczaj osoby dobierane do takich prób mają jakieś doświadczenia związane z celem badania.

Ad. 2 Probabilistyczny

Dobór losowy prosty ma na celu stworzenie warunków, w których każdy element będzie miał jednakową szansę na znalezienie się w próbie

Systematyczne dobieranie próby $k=N/n$ (losujemy tylko jedną osobę, a potem np. co ósmą

Reprezentatywność próby można czasami zwiększyć, posługując się losowym doбором wartwowym. Populacja, z której dobieramy próbę, jest dzielona na podzbiory nazywane warstwami. Próbę dobieramy z każdej warstwy osobno.

Metody badań ankietowych

1. Sondaże pocztowe

Choć roga pocztowa jest szybka i wygodna, sondaże pocztowe narażone są na tendencyjność, ponieważ nie wszyscy respondenci odsyłają ankiety z powrotem. Ostateczna próba pozyskania w sondażu może nie być reprezentatywna dla badanej populacji.

2. Bezpośrednie

Sondaże bezpośrednie są kosztowne, jednak umożliwiają badaczom większą kontrolę nad sposobem wypełniania kwestionariuszy.

Jeśli ankieterzy naprowadzają respondentów na odpowiedzi lub niedokładnie je zapisują, mówimy o tendencyjności ankietera.

3. Telefoniczne

Pomimo wad, metoda ta jest wykorzystywana do prowadzenia krótkich sondaży. Do wad należy obecność dystraktorów, przeszkadzamy osobie do której dzwonic. Respondentami mogą być osoby posiadające telefon, istnieje ryzyko tendencyjności doboru. Metoda ta nie eliminuje również problemu tendencyjności ankietera. Istnieje także limit czasu, podczas którego respondenci chcą rozmawiać przez telefon, a słysząc w nim bezosobowy dla nich głos mogą odpowiadać inaczej

4. Internetowe

Badania internetowe mają podobne wady jak sondaże telefoniczne. Nie można dotrzeć do badanych z wszystkich grup.

Plan oparty na porównaniach poprzecznych

Prowadząc badanie oparte na porównaniach poprzecznych, naukowcy dobierają z populacji jedną lub więcej prób w tym samym czasie.

Plan oparty na porównaniach poprzecznych pozwala badaczom na opis właściwości populacji lub różnic między dwoma – lub więcej – populacjami.

Na podstawie wyników sondażu można przewidywać dalsze efekty.

Plan sukcesywnych prób niezależnych

W planie sukcesywnych prób niezależnych różne próby z populacji wypełniają sondaż w kolejnych odstępach czasu.

Plan prób niezależnych nie pozwala badaczom wnioskować jak indywidualni respondenci zmieniają się w czasie. W badaniach opartych na planie sukcesywnych prób niezależnych spotykamy problem, gdy próby dobrane z populacji nie są w podobnym stopniu reprezentatywne dla populacji, czyli kiedy próby nie są porównywalne.

Plan badań podłużnych

W planie badań podłużnych Ci sami respondenci są ankietowani wielokrotnie w celu określenia zmian, które wystąpiły wśród indywidualnych respondentów.

Ze względu na korelacyjną naturę wyników badań ankietowych, trudno jest określić przyczyny zmian pojawiających się u tych samych osób badanych w różnych momentach.

W wyniku zmniejszania się liczby osób badanych w miarę upływu czasu („śmiertelność” respondentów) ostateczna próba może nie być już reprezentatywna dla populacji.

Fakt, że respondenci wypełniają kwestionariusz więcej niż jeden raz, może wpływać na ich odpowiedzi.

Zgodność pomiędzy zachowaniem deklarowanym w faktycznym

Badania ankietowe posługują się pomiarem reaktywnym, ponieważ respondenci wiedzą, że ich odpowiedzi są zapisywane.

Dążenie do aprobaty społecznej może powodować, że respondenci nie będą odpowiadać zgodnie z prawdą, ale zgodnie z wyobrażeniami o tym, co powinni odpowiedzieć.

Badacze mogą szacować prawidłowość odpowiedzi w sondażach, porównując ich wyniki z danymi archiwalnym lub obserwacjami zachowania.

18.04 Wykład dr Januszek

Plan eksperymentalny

A. Charakterystyka planu eksperymentalnego

Celem eksperymentu jest ustalenie przyczyn danego zjawiska -> kanony indukcji Milla

Kanon jednej różnicy: „Jeżeli przypadek, w którym dane zjawisko zachodzi, oraz przypadek, w którym ono nie zachodzi, mają wszelkie okoliczności wspólne, wyjąwszy jedną, i przy tym ta jedna zachodzi tylko w przypadku pierwszym, to okoliczność, co do której jedynie te dwa przypadki się różnią, jest skutkiem albo przyczyną, albo nieodzowną częścią przyczyny danego zjawiska”

Eksperyment to taki sposób sprawdzania zależności przyczynowej, w której manipulujemy zmienną niezależną i sprawdzamy efekty zmiennej zależnej starając się kontrolować inne czynniki, które mogłyby wpłynąć. Jeśli tylko manipulujemy tą jedną zmienną niezależną, to różnice w zmiennej zależnej będą nam mówić o tej zależności przyczynowej.

$X, A, B, C \rightarrow Y$

$\sim X, A, B, C \rightarrow \sim Y$ stąd $X \rightarrow Y$

DEFINICJA: Model eksperymentalny jest to model sprawdzania hipotezy o zależnościach między zmienną (zmiennymi) zależną i zmienną (zmiennymi) niezależną-główną, który zakłada:

- manipulację co najmniej jedną zmienną niezależną główną,
- kontrolowanie pozostałych zmiennych ubocznych i zakłócających
- dokonywanie pomiaru zmienności zmiennej (zmiennych) zależnej pod wpływem oddziaływania zmiennych niezależnych.

Warunki wnioskowania przyczynowo skutkowego:

- współzmiennność,
- porządek czasowy,
- eliminacja alternatywnych wyjaśnień.

Cechy dobrego eksperymentu:

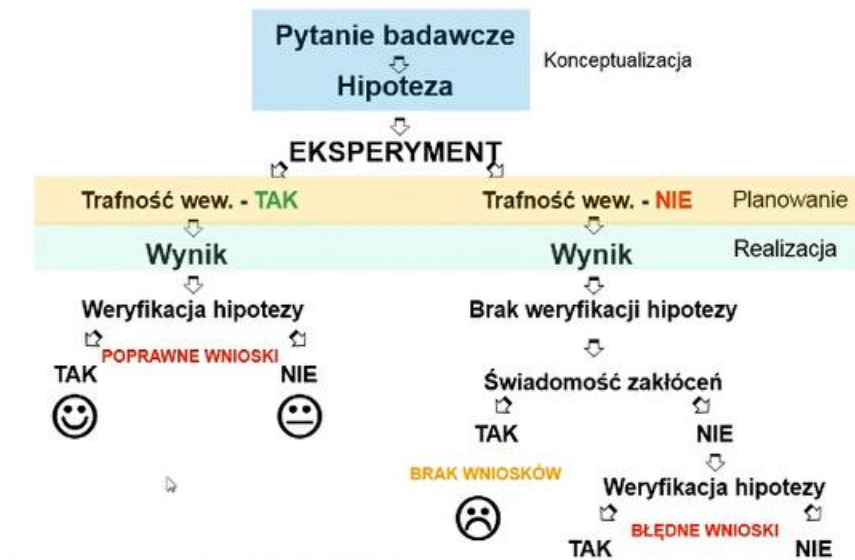
1. **trafność wewnętrzna – najważniejsze pojęcie badania eksperymentalnego, bez niego brak prawidłowych wniosków.**
2. wysoka trafność zewnętrzna,
3. Wysoka rzetelność (powtarzalność/replikowalność).
4. Wysoka wrażliwość = zdolność do wykrywania słabych zależności

Pojęcie trafności wewnętrznej

metafora hydrauliczna - badanie nietrafne wewnętrznie



Eksperyment musi być szczelny, jeśli nie mam alternatywnych wyjaśnień to otrzymuję konkretny wynik rozstrzygający



Czynniki zakłócające trafność wewnętrzną.

Trafność wewnętrzna – czy X rzeczywiście wpływa na Y?

1. Plan musi być adekwatny do danej hipotezy (np. zależność liniowa vs. Krzywoliniowa). Planowanie!
Jeśli zależność jest inna niż liniowa, to nie powinnam tworzyć planów tylko z jedną grupą eksperymentalną.
2. Efektywna manipulacja – maksymalizacja wariancji Y ze względu na postępowanie eksperymentalne
3. Badacz musi kontrolować WSZYSTKIE czynniki mogące mieć znaczenie dla Y – czynniki zakłócające, np. pogoda w trakcie badań, bo ciśnienie atmosf. Wpływa na samopoczucie.

Nastawienie badacza na potwierdzenie hipotezy badawczej:

negatywne oczekiwania -> negatywne efekty (efekt Golema)

pozytywne oczekiwania -> pozytywne efekty (efekt Galatei)

Status motywacyjny osoby badanej

- motywacja negatywna: niechęć do współpracy, brak dostatecznej uwagi, utrudnianie, odmawianie dalszej współpracy
- motywacja pozytywna: nadgorliwość, zgadywanie "hipotezy", symulowanie oddziaływania eksperymentalnego

Wykład 30.05 dr Januszek – analiza częstości

Statystyki			
mat_20			
1.00	N	Ważne	273
		Braki danych	0
	Średnia		59.8828
	Mediana		59.0000
	Dominanta		50.00
	Rozstęp		82.00
	Minimum		16.00
	Maksimum		98.00
2.00	N	Ważne	289
		Braki danych	0
	Średnia		59.3495
	Mediana		59.0000
	Dominanta		50.00
	Rozstęp		85.00
	Minimum		11.00
	Maksimum		96.00
3.00	N	Ważne	300
		Braki danych	0
	Średnia		59.5467
	Mediana		59.0000
	Dominanta		53.00
	Rozstęp		81.00
	Minimum		19.00
	Maksimum		100.00

Jakie są wnioski tej analizy? Czy te rejony różnią się wynikami matur w 2020? Brak istotnych różnic.

Miary tendencji centralnych i miar rozproszenia.

Odchylenie standardowe w percentylach: percentyle; mediana jest szczególnym percentylem, możemy zredukować informację do prostszych miar.

Rozkłady zmiennej, rozkład normalny

Rozkład normalny wzór

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

Tabele frekwencji:

- Służą do przedstawienia frekwencji występowania określonych wartości zmiennej w badanej grupie,
- rozkład częstości można utworzyć dla dowolnej zmiennej mierzonej na dowolnej skali,
- w przypadku, gdy zmienna ma wiele wartości (duży zakres wyników), tworzy się wyniki pogrupowane w tzw. Przedziały klasowe
- grupowanie wyników wiąże się z utratą informacji (błąd grupowania)
- tabele frekwencji można uzupełnić o kolumny zawierające liczebności względne (np., %) i skumulowane
- alternatywą dla rozkładu frekwencji może być diagram „pień i liść”

Zasady, zalecenia przy tworzeniu przedziałów klasowych

- a. rozłączność przedziałów
- b. równa szerokość wszystkich przedziałów
- c. ciągłość przedziałów
- d. na początku tabeli wyniki najwyższe,
- e. 10-20 przedziałów
- f. szerokość przedziału 2,3, 4, 5, 10,25, 50
- g. dla potrzeb wykresu – wartości nieparzyste szerokości przedziału
- h. dolna granica przedziału = wielokrotność szerokości

Kwantyle

Wartości cechy badanej zbiorowości, które dzielą ją na określone części pod względem liczby jednostek. Szeregi, z których wyznaczane są kwantyle, muszą być uporządkowane wg rosnących (lub malejących) wartości cechy.

Do najczęściej używanych kwantyli zaliczamy:

- medianę (podział na dwie równe części po 50 %)
- kwartyle (podział na 4 równe części po 25 %)
- decyle (podział na 10 równych części po 10%)
- percentyle (podział na 100 równych części po 1%)

Centyle i rangi centylowe

centyl – punkt na skali pomiarowej, poniżej którego znajduje się określony procent obserwacji w rozkładzie (dowolne wartości)

ranga centylowa – procent obserwacji w rozkładzie, które znajdują się poniżej danego punktu na skali pomiarowej (wartości od 0 do 100)

Miary tendencji centralnej

Średnią liczymy wyłącznie dla skal interwałowych lub stosunkowych.

Cechy średniej:

1. Każda zmienna mierzona na skali przedziałowej lub ilorazowej ma swoją średnią (nie każda ma np. dominantę)
2. Średnia jest liczona na podstawie wszystkich wartości zmiennej
3. Suma odchyleń wszystkich pomiarów w zbiorze od ich średniej arytmetycznej jest równa 0
4. Suma kwadratów odchyleń od średniej arytmetycznej jest mniejsza niż suma kwadratów odchyleń od dowolnej innej wartości
5. Średnia jest bardziej stabilna od innych miar tendencji centralnej w przypadku losowania prób z populacji
6. Na wielkość średniej wpływają wyniki odstające i skrajne.

Średnia ważona.

Przykład: średnia klasy A ($n_A=20$) z testu z historii wynosi 72,4 pkt; średnia klasy B ($n_B=60$ pkt). Ile wynosi średnia obu klas?

Rozwiązanie:

$$M_{(A+B)} = (M_A + M_B) / 2 = 66,2 \text{ zł}$$

$$M_w = (M_A + n_A + M_B + n_B) / (n_A + n_B) = 64,96$$

Średnia harmoniczna

Średnią harmoniczną liczymy wyłącznie dla skali stosunkowej. Stosuje się ją w sytuacji, gdy wartości zmiennej są wyrażone w jednostkach względnych (np. km/h, osoby na km²)

Zadanie: Gęstość zaludnienia w dwu 60-tysięcznych miastach wynosi odpowiednio: 400 osób/km² i 600 osób/km². Jaka jest przeciętna gęstość zaludnienia?

Rozwiązanie 1:

$$M = (M_A + M_B) / 2 = (400 + 600) / 2 = 500 \text{ osób/km}^2 \leftarrow \text{ŹLE!}$$

$$HM = 2 / (1/x_1 + 1/x_2) = 480 \text{ osób/km}^2 \leftarrow \text{DOBRZE}$$

Średnia geometryczna

Liczymy wyłącznie dla skali stosunkowej

Średnia geometryczna jest stosowana do opisu względnych zmian wartości, tj. średniego tempa zmian.

Mediana

Mediana (wartość środkowa); Me – środkowa liczba w uporządkowanej niemalejąco próbce (dla próbki o liczności nieparzystej) lub średnią arytmetyczną dwóch liczb środkowych (dla próbki o liczności parzystej).

Medianę liczymy dla skali porządkowej lub wyższej.

Zadanie:

Oblicz medianę dla następujących wyników:

A: 9, 4, 5, 3, 10, 1, 0

B: 2, 5, 3, 1, 4, 8

Rozwiązanie:

ad A.

- sortujemy wyniki: 0, 1, 3, 4, 5, 9, 10
- liczba pomiarów = 7 (nieparzysta)
- $Me_A = 4$ (wynik 4 osoby)

ad B.

- sortujemy wyniki: 1, 2, 3, 4, 5, 8
- liczba pomiarów = 6 (parzysta)
- $Me_B = (3 + 4) / 2 = 3,5$ (średni wynik 3 i 4 osoby)

Modalna (dominanta)

Wartością modalną (modą, dominantą) próby o powtarzających się wartościach nazywamy najczęściej powtarzającą się wartość.

Modalną możemy liczyć dla wszystkich skal

Miary rozproszenia

Rozstęp, minimum i maksimum

$$R = x_{max} - x_{min}$$

gdzie:

x_{max} – wynik maksymalny

x_{min} – wynik minimalny,

Rozstęp ćwiartkowy (rozstęp kwartylny)

$$R_q = Q_3 - Q_1$$

gdzie:

R_q – rozstęp ćwiartkowy

Q_3 – trzeci kwartyl,

Q_1 – pierwszy kwartyl

Odchylenie przeciętne od średniej

$$OP_M = \frac{1}{n} \sum_{i=1}^n |x_i - M|$$

gdzie:

OP_M – odchylenie przeciętne od średniej

M – średnia,

n – liczebność grupy,

x_i – wynik i-tej jednostki.

Wariancja i odchylenie standardowe

a) wariancja

$$SD^2 = \frac{1}{n} \sum_{i=1}^n (x_i - M)^2$$

gdzie:

SD^2 – wariancja

M – średnia,

n – liczebność grupy,

x_i – wynik i-tej jednostki.

b) odchylenie standardowe

$$SD = \sqrt{SD^2}$$

Rozkłady zmiennej, rozkład normalny

Rozkład normalny – cechy:

- kształt krzywej normalnej zależy od średniej arytmetycznej, odchylenia standardowego; wartości μ determinuje położenie wartości maksymalnej na osi zmiennej niezależnej, zaś wartość σ określa stopień spłaszczenia krzywej
- powierzchnia pod krzywą rozkładu standaryzowanego jest równa jedności,
- rozkład jest symetryczny o najwyższej wartości Y dla pomiaru X równego wartości średniej μ ,
- zmienna X może przyjmować wartości z przedziału $(-\infty, +\infty)$,
- wszystkie wartości Y są dodatnie
- około 68,26% powierzchni pod krzywą zawarte jest w przedziale jednego odchylenia standardowego od wartości średniej,
- około 95,46% powierzchni pod krzywą zawarte jest w przedziale dwóch odchylen standardowych od

wartości średniej

- około 99,73% powierzchni pod krzywą zawarte jest w przedziale trzech odchyłeń standardowych od wartości średniej

Graficzna prezentacja statystyk opisowych

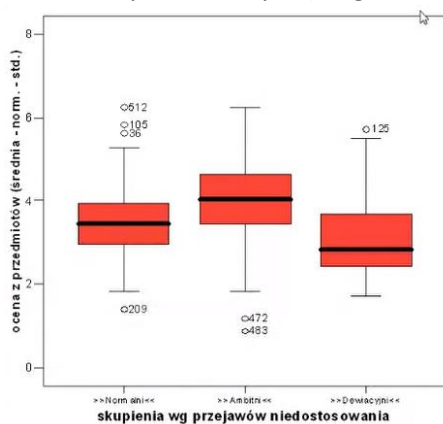
Wykres skrzynkowy:

gruba, czarna linia pośrodku skrzynki wyznacza medianę,

górną i dolną skrzynki to wartości pierwszego i trzeciego kwartyla (odległość między tymi dwiema wartościami to tzw. Rozstęp ćwiartkowy),

wąskie pionowe linie zakończone liniami poziomymi zawierają te obserwacje, które leżą w zakresie 1,5 rozstępu ćwiartkowego od skrzynki

kółka i gwiazdki to obserwacje nietypowe: kółka to przypadki odstające (1,5 do 3 rozstępów ćwiartkowych od skrzynki), a gwiazdki to przypadki skrajne (ponad 3 rozstępy ćwiartkowe od skrzynki)



5.06 dr Grabowski – wnioskowanie statystyczne (testy parametryczne/nieparametryczne)

Wnioskowanie statystyczne to dział statystyki zajmujący się problemami uogólniania wyników badania próby losowej na całą populację.

Szacowanie błędów dotyczy badania statystycznego. Badanie statystyczne różni się od opisu statystycznego. Dwie grupy metod definiują dwa działy wnioskowania statystycznego – estymacja (szacowanie wartości nieznanymi parametrów rozkładu) i weryfikacja hipotez statystycznych.

Estymacja dotyczy szacowania przedziałów ufności (duże znaczenie! Ale nie wiem w czym:))

Wnioskowanie częstościowe – oparte na kontroli błędów decyzyjnych, perspektywa tradycyjna

Wnioskowanie Bayesowskie – oparte na aktualizacji prawdopodobieństwa subiektywnego w świetle napływających danych; dwie hipotezy

Wnioskowanie = etapy

1. Ustalanie schematu badania i liczby oraz sposobu pomiaru zmiennych zależnych i niezależnych – jest to ściśle powiązane ze stawianymi pytaniami i hipotezami badawczymi
2. Formułowanie hipotezy zerowej – wiąże się automatycznie z wyborem testu statystycznego
3. Sprawdzanie założeń wybranej techniki statystycznej (dla testu t: czy zmienna zależna jest badana na skali ilościowej? Jeżeli byłoby to badanie kwestionariuszem to jest to ilościowa; czy zmienna posiada rozkład zgodny z rozkładem normalnym? Czy wariancja w obu grupach jest taka sama? – test Levene'a;

jeżeli test jest nieistotny to nie możemy odrzucić hipotezy zerowej, a zatem założenie zostało spełnione – wariancje są sobie równe; czy grupy są równoliczne?

4. Obliczenie wartości testu według ustalonej procedury (wzroy) – na szczęście robi to za nas program statystyczny, taki jak np. SPSS

5. Odczytywanie wartości krytycznych dla danej statystyki z tablic statystycznych, odnoszenie otrzymanej wartości testu do ustalonego poziomu istotności – jakby strasznie to nie brzmiało, też robi to za nas program typu SPSS – np. odczytywanie wartości krytycznej dla testu t i porównywanie go do wartości przeliczonej przeze mnie. Jeżeli wartość którą otrzymałem była większa od tej krytycznej odczuwało się hipotezę zerową. Jeżeli prawdopodobieństwo będzie wynosić np. 0.01 to będzie to oznaczać, że jest niższe od klasycznej wartości prawdopodobieństwa, które wynosi 0.05 zatem możemy na poziomie 0.05 odrzucić hipotezę zerową

6. Podjęcie decyzji o odrzuceniu bądź nie hipotezy zerowej. Gdy odrzucamy hipotezę zerową, przyjmujemy hipotezę alternatywną. Jeżeli nie odrzucamy hipotezy zerowej to po prostu nie mamy prawa jej odrzucić. Nie oznacza to, w sensie statystycznym, że ją przyjmujemy, tylko na podstawie tego badania, które wykonaliśmy nie mamy prawa odrzucić hipotezy zerowej.

7. Interpretacja wyników, najczęściej na podstawie dodatkowych statystyk opisowych

Określenie hipotezy zerowej i alternatywnej.

Zgromadzenie rzetelnych danych za pomocą dobrze zaprojektowanego badania.

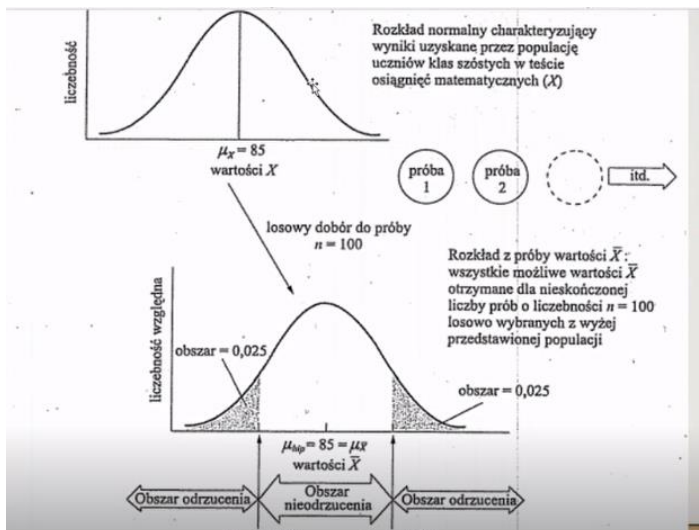
Obliczanie statystyk testową na podstawie danych. Bazujemy na rozkładzie testu t studenta. Stopnie swobody dotyczą liczebności osób badanych i zmiennych. Ten rozkład jest podstawą, bo w jego obszarze wyróżniamy obszar nieodrzczenia i odrzczenia hipotezy zerowej.

Ustalenie wartości P dla statystyki testowej.

Określenie na podstawie wartości P czy należy odrzucić H_0 czy nie

Świadomość, że wniosek może być nieprawidłowy po prostu z przyczyn losowych.

Stawianie hipotezy alternatywnej kierunkowej. Znak wyniku testu, który otrzymujemy. Znak ujemny – mówimy o sytuacji, że średnia jest wartością niższą, jeżeli dodatni – średnia jest wartością wyższą. Tak wzbogadzamy hipotezę alternatywną.



Hipotezą zerową, oznaczaną przez H_0 jest hipoteza o wartości jednego (lub ielu)parametru populacji.

Tę hipotezę traktujemy jako prawdziwą, dopóki nie uzyskamy informacji statystycznych dostatecznych do zmiany naszego stanowiska.

Hipotezą alternatywną, oznaczana przez H_1 jest hipoteza przypisująca parametrowi (parametrom) populacji wartość niezgodną z przypisaną mu (im) przez hipotezę zerową.

Hipoteza zerowa i alternatywna tworzą pewną parę hipotez dopełniających się, które uwzględniają wszystkie możliwe wartości parametru lub parametrów. Przykładem takich hipotez może być para:

$H_0: \mu = 100$,

$H_1: \mu \neq 100$.

Hipoteza zerowa zakłada, że średnia w populacji jest równa 100, a hipoteza alternatywna, że jest różna od 100. Jedna i tylko jedna z tych dwóch hipotez jest prawdziwa: albo średnia w populacji jest równa 100 albo też jest równa którejś z nieskończenie wielu liczb różnych od 100.

Często hipoteza zerowa opisuje sytuację, która istniała dotąd lub jest wyrazem pewnego naszego przekonania, które chcemy sprawdzić, żeby wiedzieć czy tego przekonania nie należy odrzucić na korzyść hipotezy alternatywnej. Sprawdzenia dokonuje się korzystając z informacji zawartej w próbie losowej. PRzy sprawdzaniu hipotezy o wartości średniej w popoulacji, a więc w sytuacji takiej, jaką opisuje wzór, sprawdzianem, czyli statystyką testu, jest średnia z próby \bar{X} -> sprawdzian hipotezy zerowej.

Sprawdzianem lub statystyką tesstu nazywamy statystykę z próby, której wartość obliczona na podstawie wyników obserwacji jest wykorzystywana do ustalenia czy możemy hipotezę zerową odrzucić, czy jej odrzucić nie możemy.

Przykładem reguły decyzyjnej zastosowanej do sprawdzenia hipotezy zerowej wskazanej we wzorze, może być reguła: „odrzuć hipotezę zerową wtedy i tylko wtedy, gdy okaże się ni mniejsze od 105 lub nie większe od 95”.

Jak patrzymy sobie na tabelę z testem t to patrzymy na prawdopodobieństwo, jeśli to p jest poniżej 0.05 to odrzucamy hipotezę zerową. Jeśli jest powyżej, to nie mamy prawa jej odrzucić. Kiedy tak się dzieje? Wówczas gdy weryfikujemy założenia poszczególnych testów, np. testu t.

W teście Levene'a oczekujemy, że wartość p będzie wyższa niż 0.05 – w teście t oczekujemy czegoś zupełnie innego! WERYFIKUJEMY HIPOTEZĘ ZEROWĄ, A POD TYM KRYJĄ SIĘ RÓŻN TREŚCI

Tablica 7.1. Stan rzeczy, decyzje i dwa rodzaje błędów

		Stany rzeczy	
		H_0	H_1
Decyzje	H_0	Śluszna decyzja	Błąd rodzaju II
	H_1	Błąd rodzaju I	Śluszna decyzja

Odrzucenie prawdziwej hipotezy zerowej

Nieodrzućenie fałszywej hipotezy zerowej

Prawdopodobieństwo popełnienia błędu pierwszego rodzaju oznaczamy przez α , a prawdopodobieństwo popełnienia błędu drugiego rodzaju przez β .

Przy wnioskowaniu statystycznym wyróżniamy 4 błędy: 1,2,3,4 rodzaju. Nie zajmujemy się tymi ostatnimi. Skupiamy się na pierwszych 2. Info dt reszty na blogu statystycznym.

Jeśli odrzucamy hipotezę zerową, a ona jest prawdziwa, to popełniamy błąd pierwszego rodzaju. Gdy odrzucamy hipotezę nieprawdziwą, to podejmujemy słuszną decyzję. Drugi błąd to nieodrzućenie fałszywej hipotezy zerowej. – ale skąd wiedzieć kiedy hipotezy są prawdziwe, a kiedy nie? I kiedy je odrzucić?

Ryzyko (prawdopodobieństwo) popełnienia błędu I rodzaju nazywane jest poziomem istotności – w badaniach staramy się zminimalizować ten błąd. W naukach społecznych za dopuszczalne przyjmuje się 5% ryzyko popełnienia tego błędu (zapisywane jako $p < 0.05$). Minimalizując ryzyko popełnienia jednego z tych błędów, zwiększamy prawdopodobieństwo popełnienia drugiego.

Oba prawdopodobieństwa dt. I i II błędu są prawdopodobieństwami względnymi, prawd. I błędu jest prawdopodobieństwem, że po pobraniu próbu i obliczeniu wartości sprawdzianu odrzucimy hipotezę zerową przy założeniu, że ta hipoteza jest prawdziwa. Natomiast prawdopodobieństwo II błędu jest prawdopodobieństwem, że po pobraniu próby i obliczeniu wartości sprawdzianu nie odrzucimy hipotezy zerowej, przy założeniu, że ta hipoteza jest fałszywa. Zachodzą więc następujące

$$\alpha = P(H_0 \text{ odrzucona} | H_0 \text{ jest prawdziwa}), \quad (7.2)$$

$$\beta = P(H_0 \text{ nieodrzucona} | H_0 \text{ jest fałszywa}). \quad (7.3)$$

równości:

Zwykle będziemy mieli do czynienia z hipotezą statystyczną, którą próbujemy odrzucić. Przed przeprowadzeniem testu znamy prawdopodobieństwo popełnienia błędu I rodzaju. To prawdopodobieństwo jest z góry ustaloną małą liczbą, powiedzmy 0,05. Świadomość, że istnieje małe prawdopodobieństwo popełnienia błędu I rodzaju, czyli odrzucenia hipotezy zerowej, gdy nie powinna być ona odrzucona, czyni odrzucenie hipotezy zerowej wnioskiem stanowczym. Nie można tego samego powiedzieć o akceptowaniu (ściślej: nieodrzućeniu) hipotezy zerowej. Jest tak dlatego że w przeciwieństwie do prawdopodobieństwa alfa, prawdopodobieństwo beta nieodrzućenia hipotezy zerowej, gdy powinna być odrzuconą nie jest na ogół z góry ustaloną, małą liczbą. Nieodrzućenie hipotezy zerowej jest zwykle niestanowczym, słabym wnioskiem. Gdy odrzucamy hipotezę zerową, jesteśmy silnie przekonani, że hipoteza powinna być odrzucona.

Reguła decyzyjna przy testowaniu hipotezy statystycznej polega na porównaniu wartości sprawdzianu z wartościami rozgraniczającymi obszary odrzucenia i nieodrzućenia (przyjęcia). Hipotezę zerową odrzucamy wtedy i tylko wtedy gdy sprawdzian wpada w obszar odrzucenia przy przyjętym poziomie istotności alfa.

Obszary odrzucenia można odnosić do określonych rozkładów, statystyk, które wyliczymy jako sprawdziany hipotezy zerowej.

Obszarem odrzucenia hipotezy statystycznej jest taki zbiór liczb, że jeżeli sprawdzian przyjmie wartość z tego zbioru, to hipotezę zerową odrzucimy. Obszar odrzucenia nazywa się też **obszarem krytycznym**. Obszar krytyczny wyznaczają **punkty (wartości) krytyczne**. Obszar krytyczny ustalany jest tak, by przed pobraniem próby prawdopodobieństwo α , że sprawdzian znajdzie się w tym obszarze, przy założeniu, że hipoteza zerowa jest prawdziwa, było równe α .

Obszarem nieodrzućenia (przyjęcia) hipotezy statystycznej (też wyznaczonym przez punkty krytyczne) jest taki zbiór liczb, że jeżeli sprawdzian przyjmie wartość z tego zbioru, to hipotezę zerową nie odrzucimy. Obszar nieodrzućenia (przyjęcia) jest wyznaczony tak, by przed pobraniem próby prawdopodobieństwo, że sprawdzian znajdzie się w tym obszarze przy założeniu, że hipoteza zerowa jest prawdziwa, było równe $1 - \alpha$.

Statystyka t. Stosujemy ją wówczas, gdy odchylenie standardowe nie jest znane.

Mamy trzy testy t:

1. dla jednej próby – porównujemy średnie w danej próbie ze średnią hipotetyczną (kryterialną) – musimy wpisać jej wartość do okienka; założenie testu: rozkład badanej zmiennej (zależnej) musi być zgodny z rozkładem normalnym (!) – jeżeli tak nie jest to zastępujemy test t testem Wilcoxa (test z grupy nieparametrycznych); jedna próba, jedna zmienna
2. Dla prób niezależnych – gdy porównujemy dwie grupy zbudowane z innych osób (grupa kontrolna i eksperymentalna; w warunkach naturalnych: kobiety vs mężczyźni); zmienna grupująca to np. płeć; w zakres zmiennej grupującej wprowadzamy zmienną dychotomiczną (zmienną dwuwartościową!)

Pytanie: mówimy, że zmienną grupującą ma być zm. Dychotomiczna. W przypadku płci to jest oczywiste, a co gdy mamy dwie grupy – kontrolną i eksperymentalną?

Po prostu musimy je oznaczyć

► Test T (T-TEST)

Statystyki dla grup				
	płeć	N	Średnia	Odchylenie standardowe
NEU	kobieta	227	24.73	7.771
	mężczyzna	178	20.47	8.284
				Błąd standardowy średniej
				.516
				.621

Test dla prób niezależnych									
Test Levene'a jednorodności wariancji				Test t równości średnich					
		F	Istotność	t	df	Istotność (dwustronna)	Różnica średnich	Błąd standardowy różnicy	95% przedział ufności dla różnicy średnich
									Dolna granica
NEU	Założono równość wariancji	2.320	.128	5.325	403	.000	4.265	.801	2.690
	Nie założono równości wariancji			5.284	368.240	.000	4.265	.807	2.678
									Górna granica
									5.840
									5.852

test F – zestaw hipotez zerowej i alternatywnej, który sprawdza założenie dotyczące testu t, które mówi o tym, że wariancje obu grup są sobie równe (wariancje powinny być do siebie podobne); w sensie statystycznym jest to postawienie H_0 = wariancja grupy pierwszej jest równa wariancji grupy drugiej

Czy możemy dla $F=2,320$ odrzucić hipotezę zerową? Widać wyraźnie, że nie, ponieważ mamy istotność powyżej 0,1, grubo powyżej 0,05, więc nie mamy prawa by jej odrzucić. Musimy spojrzeć na wiersz

testu t dotyczący „założono równość wariancji” (drugi wiersz dotyczący poprawki testu Welcha). Patrzymy teraz na istotność dwustronną .000 – grubo poniżej 0,05 a zatem odrzucamy hipotezę zerową mówiącą, że średnia kobiet i mężczyzn **są sobie równe** i przyjmujemy hipotezę alternatywną. t w wierszu *założono równość wariancji* jest dodatnie – oznacza to, że średnia pierwsza jest wyższa od średniej drugiej, np. średnia neurotyczność kobiet jest wyższa niż średnia neurotyczność mężczyzn.

Zazwyczaj dwie średnie z różnych od siebie grup będą się różnić. Test t Studenta powie nam jednak czy owe różnice są istotne statystycznie – czy nie są przypadkowe. Hipoteza zerowa takiego testu będzie brzmiała *H0: Średnia liczba wypalanych papierosów w grupie mężczyzn jest taka sama jak średnia liczba wypalanych papierosów w grupie kobiet*. [Hipoteza](#) alternatywna z kolei *H1: Kobiety będą różnić się od mężczyzn pod względem liczby wypalanych papierosów w ciągu miesiąca*.

3. Dla prób zależnych – dwa pomiary tej samej zmiennej i tej samej grupy, ta sama osoba wchodzi zarówno w skład grupy eksperymentalnej i kontrolnej; jedna obserwacja

Badam ugodowość w parach: ale jakich????? Kto to jest Uje

Statystyki dla prób zależnych					
		Średnia	N	Odchylenie standardowe	Błąd standardowy średniej
Para 1	UGD1	25.06	405	6.418	.319
	UGD2	27.47	405	7.023	.349

Korelacje dla prób zależnych			
		N	Istotność
Para 1	UGD1 & UGD2	405	.000

Test dla prób zależnych								
Różnice w próbach zależnych								
		Średnia	Odchylenie standardowe	Błąd standardowy średniej	95% przedział ufności dla różnicy średnich		t	Istotność (dwustronna)
					Dolna granica	Górna granica		
Para 1	UGD1 - UGD2	-2.412	8.493	.422	-3.242	-1.583	-5.716	.000

Kliknij zaskrobnij

Ujemne t świadczy o tym, że druga średnia jest większa od pierwszej (bo od pierwszej odejmujemy drugą wzoru). Istotność dwustronna jest poniżej 0,05, zatem odrzucamy hipotezę zerową mówiącą, że średnie obu pomiarów są sobie równe.

Testy t możemy zastąpić testami nieparametrycznymi, wtedy gdy nie są spełnione założenia wykorzystania tych testów, np. rozkłady zmiennych oddalają się od rozkładu normalnego, mamy różne wariancje, rozrzut w pierwszej grupie jest inny niż w drugiej grupie. Jeżeli w tych sytuacjach będziemy korzystać z testu to będzie on obciążony, dlatego należy korzystać z testów nieparametrycznych w tym przypadku

Założenia testów t :

- każda próba jest wybrana losowo z populacji przez nie reprezentowanych,
- dwie próby losowe muszą zostać wybrane niezależnie,
- próby są wybierane zgodnie ze schematem losowania zwrotnego,
- rozkład wartości $X - Y$ z próby jest zgodny z rozkładem normalnym

Statystyka t może być wykorzystywana do testowania hipotez dt. Średniej, przy założeniu, że rozkład zmiennej zależnej jest zgodny z rozkładem normalnym.

Już sam fakt, że testy t są stosowane do porównywania średnich, narzuca to, że zmienna zależna musi być mierzona na skali ilościowej (inaczej nie będzie możliwe obliczenie średniej).

Test t dla prób niezależnych wymaga ponadto spełnienia założenia o jednorodności wariancji w porównywanych grupach (wariancja, czyli rozproszenie wyników w porównywanych próbach, powinna być podobna). Jednorodność wariancji jest liczona za pomocą testu Levene'a. Wariancja, czyli rozproszenie wyników w porównywanych próbach powinna być podobna. W sensie statystycznym: sprawdzamy hipotezę zerową mówiącą, że wariancja próby A jest równa wariancji próby B.

Należy też pamiętać, aby porównywane grupy były równoliczne (a przynajmniej nie powinny różnić się znacznie pod względem liczby osób badanych) – weryfikujemy to z pomocą testu chi kwadrat (zobacz Statystyczny Drogowskaz cz I rozdz. 5 – przykłady i weryfikowanie równoliczności grupy)

Testy T: dla jednej próby

Porównujemy średnią z próby z pewnym kryterium.

Hipoteza zerowa: oceny porównawcze nie będą się różniły od wartości 0 $H_0: \mu=0$

Hipoteza alternatywna: oceny porównawcze będą różniły się od wartości 0 $H_1: \mu \neq 0$

Test T: dla dwóch prób zależnych

Zmienne i podstawowe założenia: zmienna niezależna wewnątrzsobowa (powtarzalny pomiar), typ ocenianej osoby jest mierzona na dwóch poziomach (Ja vs przeciętny student);

Testujemy różnice

Zmienna zależna – ocena towarzyskości, mierzona na skali ilościowej

Test dla dwóch prób zależnych

Pomiar pierwszy i drugi danej zmiennej

$H_0: M1=M2$

$H_{11}: M1 \neq M2$

Pierwsza osoba pary bliźniąt = druga osoba tej pary

Test T dla dwóch prób niezależnych

Zmienne i podstawowe założenia:

zmienna niezależna międzyobiektowa, typ ocenianej osoby jest mierzony na dwóch poziomach (ja vs przeciętny student). Jedną zmienną testowaną w obu grupach

Testujemy różnice

Zmienna zależna ocena towarzyskości, mierzona na skali ilościowej

Próba kobiet i próba mężczyzn

$H_0: M1= M2$

$H_1: M1 \neq M2$

Grupa kontrolna vs grupa eksperymentalna

Wartość t jest tym większa, a tym samym szanse na znalezienie istotnych statystycznie różnic są tym większe im:

- większa różnica średnich,

- mniejsze odchylenie standardowe
- większa próba

Im większa próba tym mniejsze różnice stają się istotne statystycznie, tym mniejsze współczynniki korelacji stają się istotne statystycznie.

Jest to zestaw pewnych zależności między wartością t , a różnicą średnich, odchyleniem standardowym i wielkością próby

Test U-Manna Whitneya – test nieparametryczny, w przypadku zmiennych nieparametrycznych, dla prób niezależnych

Jak sprawdzić czy zmienna ma rozkład normalny? Weryfikujemy rozkład zmiennej jaką jest neurotyczność w kontekście próby mężczyzn i kobiet. Eksploracja -> zmienna zależna: wrzucamy zmienną, której rozkład chcemy sprawdzać; lista grupująca: zmienna grupująca (np.. płeć)-> wykresy: zaznaczamy wykresy z normalnością rozkładu ->

Próby duże to powyżej 100 osób. Przeliczając normalność rozkładu dla takiej próby patrzemy na wynik testu Kołomogorowa-Smirnowa. Wartość istotności znalazła się poniżej wartości 0,05, a zatem musimy odrzucić hipotezę zerową mówiącą o tym, że rozkłady neurotyczności w obu próbach są zgodne z rozkładem normalnym. W tym wypadku zamiast testu t dla prób niezależnych powinniśmy wykonać test U Manna Whitneya .

Tab. 1. Podstawowe miary wielkości efektu, miary, które są interpretowane podobnie, oraz orientacyjne wielkości ułatwiające interpretację¹

Podstawowe miary wielkości efektu	Miary siły efektu, których wielkość interpretujemy w podobny sposób	Efekty słabe	Efekty średnie	Efekty silne	Efekty bardzo silne
d (Cohena)	Δ , d Glassa, g Hedgesa	0,2-0,5	0,5-0,8	0,8-1,3	> 1,3
r (Pearsona)	R, ϕ , ρ , β , τ (tau) oraz ich pochodne*	0,1-0,3	0,3-0,5	0,5-0,7	> 0,7
r^2	R^2 , η^2 , ω^2 , ε^2	0,01-0,09	0,09-0,25	0,25-0,49	> 0,49

Źródło: na podstawie J. Cohen (1988) i J.A. Rosenthal (1996).

*Pochodne tych wielkości efektów są oznaczane dodatkowymi symbolami, np. R_{adj}^2 , ω_p itp.

<https://effect-size-calculator.herokuapp.com/>

Współczynniki korelacji Pearsona. Hipoteza 0: Współczynnik korelacji Pearsona jest istotnie różny od zera.