

Wykład 9: Moc

- Moc testu to prawdopodobieństwo odrzucenia H_0 , gdy prawdziwa jest H_A
[Moc=czułość testu]
- $Moc = 1 - Pr(\text{nie odrzucamy } H_0, \text{ gdy prawdziwa jest } H_A)$
 $= 1 - Pr(\text{błąd II-go rodzaju}) = 1 - \beta$
- Na ogół chcemy, aby test miał dużą moc

- Moc zależy od:

Wielkość efektu

- Wielkość efektu = sygnał / szum = $(\mu_1 - \mu_2) / \sigma$;
- W tabelach dana moc jednostronnego testu Studenta dla dwóch niezależnych prób na poziomie istotności $\alpha=0.01$ w funkcji rozmiaru próby i *wielkości efektu*.
- Przykład: Aby mieć 90% pewności, że jednostronny test Studenta na poziomie istotności $\alpha=0.01$ wykryje różnicę między średnimi równą 0.8σ , musimy pobrać próby o rozmiarze $n_1 = n_2 =$
- Aby oszacować σ , często wykonuje się badania wstępne.

Moc jednostronnego testu Studenta dla $\mu_1 = \mu_2$ na poziomie istotności $\alpha=0.01$

d												
n	d _c	.10	.20	.30	.40	.50	.60	.70	.80	1.00	1.20	1.40
8	1.31	.02	.03	.04	.05	.08	.12	.14	.19	.30	.43	.57
9	1.22	.02	.03	.04	.06	.09	.13	.16	.22	.35	.49	.63
10	1.14	.02	.03	.04	.07	.10	.14	.18	.25	.40	.55	.70
11	1.08	.02	.03	.05	.07	.11	.15	.21	.28	.45	.61	.76
12	1.02	.02	.03	.05	.08	.12	.17	.23	.31	.49	.66	.81
13	.98	.02	.03	.05	.08	.13	.19	.26	.34	.53	.71	.85
14	.94	.02	.03	.06	.09	.14	.20	.28	.38	.57	.75	.88
15	.90	.02	.04	.06	.10	.15	.22	.31	.41	.61	.79	.90
16	.87	.02	.04	.06	.10	.16	.24	.34	.44	.64	.82	.92
17	.84	.02	.04	.07	.11	.18	.26	.36	.47	.68	.85	.94
18	.81	.02	.04	.07	.12	.19	.27	.38	.49	.71	.87	.95
19	.79	.02	.04	.07	.13	.20	.29	.40	.51	.74	.89	.96
20	.77	.02	.04	.08	.13	.21	.30	.42	.54	.76	.91	.97
21	.75	.02	.05	.08	.14	.22	.32	.44	.56	.79	.93	.98
22	.73	.02	.05	.08	.15	.23	.34	.46	.59	.81	.94	.98
23	.71	.02	.05	.09	.15	.24	.36	.48	.61	.83	.95	.99
24	.70	.02	.05	.09	.16	.25	.37	.50	.64	.85	.95	.99

Moc jednostronnego testu Studenta dla $\mu_1 = \mu_2$ na poziomie istotności $\alpha = .01$												
n	dc	d										
		.10	.20	.30	.40	.50	.60	.70	.80	1.00	1.20	1.40
25	.68	02	05	10	17	27	39	53	66	87	96	99
26	.67	02	05	10	17	28	41	55	68	89	97	99
27	.65	02	05	10	18	29	42	57	70	90	97	*
28	.64	02	05	11	19	30	44	59	72	91	98	
29	.63	02	06	11	19	31	46	60	74	92	98	
30	.62	03	06	11	20	32	48	62	75	93	99	
31	.61	03	06	12	21	34	50	64	77	94	99	
32	.60	03	06	12	22	35	51	66	79	94	99	
33	.59	03	06	13	22	36	52	67	80	95	99	
34	.58	03	06	13	23	37	53	69	81	95	99	
35	.57	03	07	13	24	38	55	70	83	96	*	
36	.56	03	07	14	25	40	56	72	84	96		
37	.55	03	07	14	26	41	58	73	85	97		
38	.55	03	07	15	26	42	60	75	86	97		
39	.54	03	07	15	27	43	61	76	87	98		
40	.53	03	07	15	28	45	62	78	88	98		
42	.52	03	08	16	30	47	64	80	90	98		
44	.51	03	08	17	31	49	67	82	91	99		
46	.49	03	08	18	33	51	69	83	93	99		
48	.48	03	08	19	34	53	71	85	94	99		

Test Wilcoxona-Manna-Whitneya

- Test Studenta wymaga normalności rozkładów w obu populacjach.
- Jak porównać dwie populacje, w których rozkład cechy nie jest normalny, a rozmiar prób nie jest na tyle duży, aby korzystać z CTG?
- Test Manna-Whitneya nie wymaga założenia o normalności.

- H_0 : w obu populacjach badana cecha ma ten sam rozkład
- H_A : badana cecha ma inny rozkład w obu populacjach
- Test wykrywa głównie różnicę położenia (np. gdy rozkłady mają różne mediany).
- Test nieczuły na różnice w parametrach rozrzutu.

Procedura:

Łączymy obie próby otrzymując zbiór $n_1 + n_2$ liczb.

Porządkujemy otrzymany zbiór:

Rangę 1 przypisujemy wartości najmniejszej,
Rangę 2 kolejnej itd.

Dla każdej próby obliczamy sumę rang wartości pochodzących z tej próby. Otrzymujemy liczby R_1 i R_2 .

Statystyka testowa to $R = \min(R_1, R_2)$.

Przykład

- Zliczamy liczbę ziaren produkowanych przez dwie odmiany pewnej rośliny.
- Dane:
- Odmiana 1: 19, 23, 25, 28, 28, 34 ($n_1 = 6$)
- Odmiana 2: 14, 18, 19, 20, 25 ($n_2 = 5$)

Test:

- Czy obie odmiany produkują przeciętnie tyle samo ziaren ?
- H_0 : Obie odmiany mają ten sam rozkład liczby ziaren
- H_A : Rozkład liczby ziaren u odmiany 2 jest inny niż u odmiany 1
- Użyjemy testu Manna-Whitneya. Tablice wartości krytycznych są zamieszczone na www

Planowanie eksperymentu

- Rodzaje badań:
 - Badania obserwacyjne
 - Badania eksperymentalne

Badania obserwacyjne

- Zbieramy informacje o istniejącej sytuacji
- Brak kontroli nad poziomem czynnika wpływu
- Dużo nieznanych powiązań. (Może się zdarzyć, że faktycznie będziemy mierzyć wpływ innego, powiązanego czynnika.)
- Niepewne wnioski naukowe.
- Przykłady:
 - Poziom rtęci u ryb z różnych jezior
 - Poziom cholesterolu u wegetarian i ``mięsożerców``
 - Czy waga noworodków zależy od tego czy matka nadużywała alkoholu?

Badania eksperymentalne:

- Wpływamy na sytuację i mierzymy wynik.
- Mamy kontrolę nad jednym lub kilkoma czynnikami (choć niekoniecznie nad wszystkimi).
- Bardziej wiarygodne wnioski naukowe
- Przykłady:
 - Obserwujemy stan pacjentów biorących lekarstwo i placebo
 - Stosujemy cztery różne nawozy i mierzymy wydajność
 - Stosujemy różne ilości dodatku do paszy dla świń i mierzymy przyrost wagi

Badania eksperymentalne cd.

- Zmienna zależna/objaśniana (odpowieź): efekt, który mierzymy
- Zmienna niezależna/objaśniająca: czynniki, które kontrolujemy, np. rodzaj lekarstwa
- Zmienne zakłócające: czynniki, nad którymi nie mamy kontroli, np. waga pacjenta
- Przypisanie – decyzja, jaki poziom czynnika zastosujemy u każdego pacjenta.

Eksperyment zrandomizowany

- Najbardziej oczywista metoda *przypisania*
- Musimy znać:
 - ✓Liczbę czynników
 - ✓Możliwe poziomy każdego czynnika
- Zadaną kombinację poziomów czynników będziemy nazywali ``zabiegiem``

Eksperyment zrandomizowany, cd.

- Wnioski wyciągamy zakładając losowe próbkowanie z populacji (próba losowa)
- Próbę rozdzielamy losowo na części, które oddane zostaną różnym „zabiegom”.

Przykład 1– testowanie lekarstwa

- 500 pacjentek zgodziło się na przetestowanie nowego lekarstwa na raka piersi
- Ta próba reprezentuje populację kobiet z rakiem piersi, po zabiegu
- Dzielimy tę próbę LOSOWO na dwie, np. równe, grupy (po 250 osób)
- Można użyć tablicy liczb pseudolosowych

Przykład 1 cd.

- Jedna grupa dostaje lekarstwo, a druga placebo
- W okresie 5-lat obserwujemy częstość nawrotu raka w obu grupach
- Jeżeli zaobserwujemy istotną różnicę w częstościach, to z dużą pewnością będziemy mogli twierdzić, że jest to wpływ lekarstwa

Przykład 2: Różne dawki

- Jeden czynnik, ale na kilku poziomach
- Lekarstwo w dawkach 0, 10, 20, 30, 40 mg
- Dzielimy pacjentki LOSOWO na 5 grup (niekoniecznie równe rozmiary)

Przykład 3: Trzy czynniki

- Cztery różne lekarstwa
- Dwa różne zabiegi chirurgiczne
- Naświetlania lub nie
- $4 \times 2 \times 2 = 16$ możliwych ``zabiegów''
- Dzielimy 500 losowo na 16 grup zabiegowych, niekoniecznie równych rozmiarów
- Im więcej czynników, tym liczniejsza powinna być próba