

Optymalna liczebność próby

Doświadczalnictwo

Autor Wanda Olech
Katedra Genetyki i Ochrony Zwierząt

Planowanie doświadczenia

- to najważniejsza część, gdyż *a priori* musimy uzyskać odpowiedź na pytania, co, kiedy, jak i na jakim materiale
- Optymalna liczebność obiektów w doświadczeniu = najmniejsza liczebność pozwalająca w odpowiednich warunkach (przy założonej istotności) zweryfikować hipotezę

Optymalna liczebność próby jest powszechnym problemem szczególnie w doświadczeniach na zwierzętach. Ma to być taka liczebność aby pozwoliła „z pozytywnym skutkiem” zweryfikować hipotezę. Ten pozytywny skutek w testach istotności to odrzucenie hipotezy zerowej. Oczywiście im liczniejsza próba to będzie łatwiej odrzucić hipotezę zerową, ale duża próba to większe koszty, a tych też chcemy unikać.

Planowanie doświadczenia

- ◉ Cele
- ◉ Rodzaje badanych cech, zdefiniowanie istniejących i wprowadzanych czynników i ocena ich efektu
- ◉ Zwierzęta (gatunek, linia, wiek, sposób wyboru do grup itd.)
- ◉ Obiekt doświadczenia (grupa czy nie)
- ◉ Hipotezy
- ◉ Sposób analizy numerycznej

Etapy planowania doświadczenia są jak wymieniono na slajdzie. Po pierwsze musi być określony cel czyli praktycznie ujmując musimy wiedzieć jaka hipoteza będzie postawiona, czyli wpływ jakich czynników mamy zbadać. Bardzo istotne jest jakie cechy będą poddawane badaniu, czy łatwo czy trudno mierzalne, czy ciągłe czy skokowe, czy mające rozkład normalny lub nie. Potem musimy wiedzieć na jakich zwierzętach będzie prowadzone doświadczenie, czy będą one utrzymywane indywidualnie czy grupowo. Przewidujemy wszystkie hipotezy. Musimy również mieć jako takie pojęcie jakich metod numerycznych, statystycznych będzie używać. Jednym słowem jest planowanie nieznanej nam przyszłości w sposób bardzo szczegółowy. Aby taki plan był możliwy przed planowaniem musimy mieć wiedzę na temat problemu, przeczytać wszystko co zostało opublikowane, co pozwoli rozpoznać problem oraz uniknąć błędów popełnianych wcześniej przez innych.

Układy i rodzaje doświadczeń

- Układy jedno i wieloczynnikowe – możliwości i konsekwencje
- Ortogonalność układu – korzyści i ograniczenia
- Rodzaje doświadczeń w efekcie wybranego układu

Czy jeden czynnik analizować, czy jednocześnie wiele czynników? Jak zwykle taka decyzja nie jest łatwa bo więcej czynników to pełniejsze doświadczenie, ale konieczne będzie większa liczebność próby, bardziej skomplikowane analizy, problemy z interpretacją interakcji. Z drugiej strony jeden czynnik to konieczność wyeliminowania wpływów innych czynników, czyli podobieństwo grup doświadczalnych. Może być wtedy mniejsza liczebność, ale czy wyrównanie grup jest możliwe i będzie powodem do zakłócenia losowości wyboru próby czyli jej reprezentatywności dla populacji.

Ortogonalność, czyli jednakowa liczebność w grupach, podgrupach daje nam korzyści. Liczebność całego doświadczenia przy podziale na jednakowo liczne grupy zawsze będzie mniejsza niż przy podziale na odbiegające od siebie pod względem liczebności grupy. Zatem jeśli możliwe dążymy do ortogonalności.

Jak zdecydujemy się na układ doświadczenia to im więcej czynników to będziemy zbliżać się do „polowego” typu doświadczeń a jak mało czynników to w pewnym stopniu środowisko musi być kontrolowane i wtedy mamy „stacjonarny” lub „laboratoryjny” typ.

Weryfikacja hipotezy

		Obiektywnie H_0	
		prawdziwa	fałszywa
decyzja	odrzucaamy	błąd I rodzaju (α)	właściwa
	przyjmujemy	właściwa	błąd II rodzaju (β)

Przypominam tabelę możliwych decyzji przy weryfikacji hipotezy. Odrzucenie hipotezy która jest obiektywnie fałszywa, czy przyjęcie prawdziwej jest w porządku – nie popełniamy błędu, ale odrzucenie prawdziwej obiektywnie hipotezy czy przyjęcie fałszywej to jest błąd. Każda decyzja jaka mamy podejmować w weryfikacji hipotezy musi być decyzją dla której znamy prawdopodobieństwo popełnienia błędu. Znać obydwa błędy jest skomplikowane i wymaga obliczenia drugie z błędów przy założeniu wartości pierwszego z nich. Z tego względu jeśli zrezygnujemy z jednej z decyzji to możemy nie obliczać drugiego błędu jedynie założyć pierwszy z nich. W taki sposób działają testy istotności. Mamy tylko decyzję odrzucić, którą podejmiemy lub nie. Oczywiście przy stosowaniu takich testów z możliwością odrzucenia hipotezy sama hipoteza musi być sformułowana w specjalny sposób. W weryfikowanych hipotezach pojawia się informacja o parametrach lub rozkładzie zmiennej losowej zawsze w specyficzny sposób czyli zawsze parametry są jednakowe, rozkłady zgodne. Odrzucenie takiej hipotezy jest celem i daje nam możliwość wyciągania wniosków, a nie odrzucenie to powrót do punktu wyjścia.

Przewidywanie liczebności doświadczenia

- Co jest obiektem doświadczenia – grupa czy pojedyncze zwierzę (planujemy liczbę obiektów)
- Jakie mogą być straty (upadki lub inne przyczyny wycofania zwierząt) w czasie trwania doświadczenia
- Znajomość problemu pozwala ocenić poziom cechy, jej zmienność, efekt zastosowanego czynnika
- „Wykonujemy” całe doświadczenie „z góry”

Jeszcze kilka uwag na temat planowania doświadczenia – my naprawdę musimy wiedzieć i przewidywać wszystkie elementy łącznie z przewidywaną reakcją zwierząt (np. śmiertelnością), łącznie z parametrami cech czyli ich zmiennością przede wszystkim. Nawet powinniśmy przewidywać jaki efekt da nam zastosowanie danego poziomu czynnika. To wykonywanie doświadczenia „a priori” nie jest łatwe, ale bez tego nie odpowiemy na pytanie jaką liczebność zaplanować.

Przykład planowania

- Dwie grupy: kontrolna i doświadczalna
- Cecha ciągła – przewidujemy średnią i odchylenie standardowe w każdej z grup
- Stawiamy i weryfikujemy hipotezę i oceniamy warunki jej odrzucenia (istotność i moc testu)

Przykład, najprostszy możliwy. Mamy jeden czynnik działający na dwóch poziomach czyli inaczej mówiąc losujemy z populacji dwie grupy i jedną nazywamy kontrolną a druga doświadczana i ta druga jest poddawana działaniu czynnika a pierwsza z nich stanowi tło. Zakładamy że nasza cecha jest cechą ciągłą i ma rozkład normalny, czyli znowu idealna sytuacja. Hipoteza w tej sytuacji jest prosta, będzie to hipoteza o równości dwóch wartości oczekiwanych.

Przebieg weryfikacji hipotezy o braku różnic między grupami

$$H_0 : m_1 = m_2$$

$$H_1 : m_1 > m_2 \text{ (lub } m_1 < m_2 \text{ lub } m_1 \neq m_2 \text{)}$$

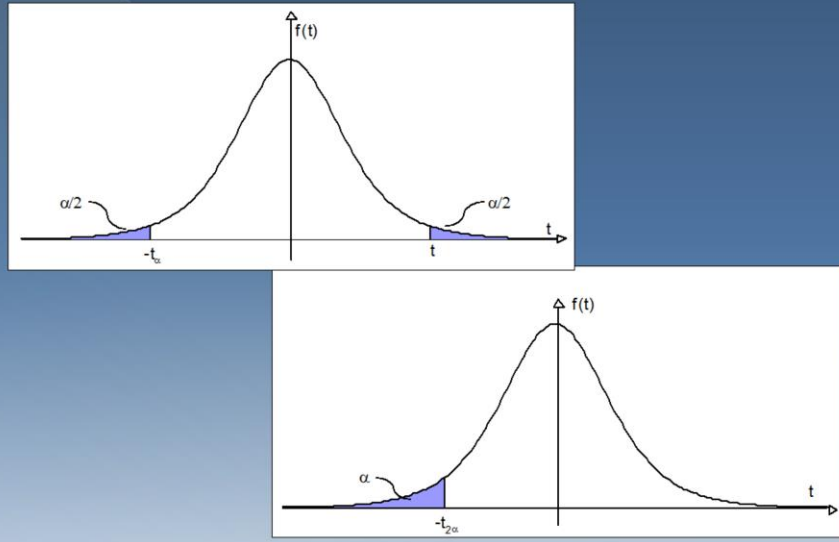
Test np. :

$$t_{emp} = \frac{\bar{x}_1 - \bar{x}_2}{S_{\bar{x}_1 - \bar{x}_2}}$$

$$S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2} \cdot \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}$$

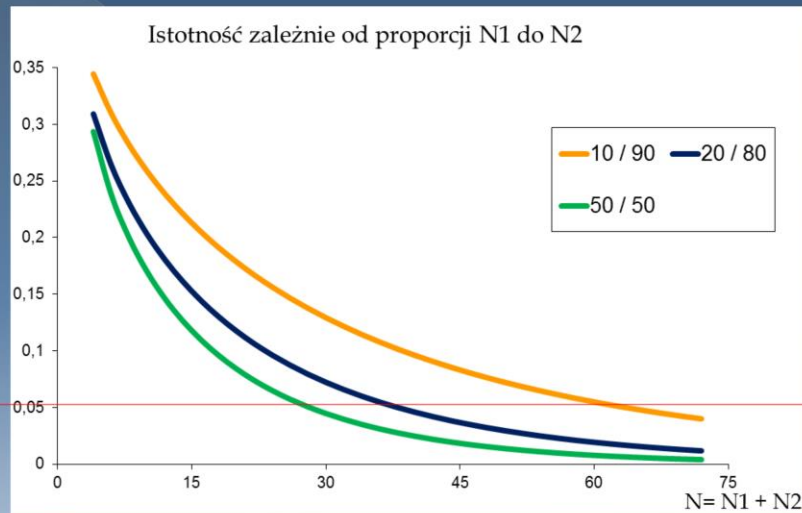
I po kolei – najpierw stawiamy hipotezę, potem dobieramy hipotezę alternatywną. I tu jest znowu możliwość decydowania, ponieważ wybór hipotezy jednostronnej (jeśli to możliwe) pozwoli odrzucić hipotezę przy mniejszej liczebności próby. Czyli warto decydować na hipotezę alternatywą prawo lub lewo- stronną. Cecha ma rozkład normalny czyli niezależnie od wielkości prób możliwe jest zastosowanie testu t-Studenta. Oczywiście są pewne ograniczenia bo test t-Studenta w tym przypadku wymaga jednorodności wariancji, ale w naszym przykładzie zakładamy że tak jest. Aby obliczyć empiryczną wartość t potrzebna jest na równica między średnimi grup i to jest nasze założenie nr oraz konieczne jest określenie błędu standardowego różnicy średnich czyli to jest nasze drugie założenie. Obydwa założone wartości wynikają z uzyskanej wiedzy na temat problemu a czasem z własnego pilotowego badania.

Obszar krytyczny



Rysunek pokazujący dwustronny lub jednostronny obszar krytyczny o takiej samej wielkości. Wartość empiryczna testu musi się znaleźć w przedziale zaznaczonym aby odrzucić hipotezę, przy hipotezie alternatywnej jednostronnej, co widać tutaj, empiryczna wartość testu może być mniejsza (pod względem wartości a nie znaku) i uda się odrzucić hipotezę. A mniejsza wartość to mniejsza liczebność.

Zalety ortogonalności



Wykres przedstawia zależność poziomu istotności (najmniejszego dla decyzji odrzucenie hipotezy) zależnie od proporcji liczebności w dwóch grupach. Zakładamy że wielkość różnicy średnich i błąd standardowy różnicy średnich i sprawdzamy istotność. Te fragmenty krzywych które są poniżej czerwonej linii dla wartości 0,05 są „satisfakcjonująca) czyli decyzja będzie odrzucenie. Najmniejsza liczebność pozwalająca na uzyskanie wymaganego poziomu istotności jest dla układu ortogonalnego, czyli podziałowi po połowie całej liczebności. Wtedy przy założonych warunkach około 26-28 obiektów w próbach łącznie wystarczy, a jak proporcja jest bardzo skośna to łączna liczebność zwiększa się. Dla przykłady podział 90:10 wymagać będzie 60 obiektów. Zatem ortogonalność to jeden z założeń pozwalających na zminimalizowanie liczebności.

Określanie liczebności

Jeśli znamy:

- Poziom istotności
- Moc testu
- Minimalną różnicę poziomu cechy między grupami (d)
- Odchylenie standardowe (S) - zakładamy jednakowe w grupach

To:

$$N = 1 + 2 \cdot C \left(\frac{S}{d} \right)^2$$

Wartości C		istotność	
		0,05	0,01
Moc testu	0,8	7,85	11,68
	0,9	10,51	14,88

Wg. Dell et al. 2002. ILAR J. 43,4:207

A teraz przewidywanie liczebności doświadczenia. Musimy założyć bardzo wiele. Poziom istotności nie jest trudny bo z reguły zakładamy 0,05 jako maksymalny w badaniach przyrodniczych. Drugim elementem jest moc testu – powinna być jak największa i z reguły do dyspozycji mamy 0,8 lub 0,9. Takie wartości są zakładane we wszystkich programach pozwalających nasz obliczać liczebność. Moc testu to 1 pomniejszone o prawdopodobieństwo błędu przyjęcia hipotezy zerowej i testy im bardziej skomplikowane mają wyższą moc. Testy nieparametryczne są słabsze niż parametryczne. Kolejne założenia to różnica między średnimi (minimalna) oraz odchylenie standardowe w grupie a założenie jest że w każdej z grup jest ono takie samo.

Dell i inni w 2002 opublikowali pracę w której dobierają stałe wartości C, które pozwalają przy założeniach obliczyć liczebność próby potrzebną do odrzucenia hipotezy.

Wielkość próby dla cech o rozkładzie zero jedynekowym

$$H_0 : p_1 = p_2$$

- W każdej grup musimy znać spodziewaną proporcję wyników „1”
- Oznaczamy przez p oraz q frekwencje w grupie kontrolnej i doświadczalnej
- Różnica frekwencji oznaczona jest przez d

Wtedy:

$$N = 0,5 \cdot d + 2 + C \left(\frac{p_k \cdot q_k - p_d \cdot q_d}{d^2} \right)$$

Trochę inaczej jest jeśli weryfikowaną hipotezą będzie hipoteza o równości prawdopodobieństw w dwóch grupach. Wór który służy ocenie potrzebnej liczebności jest jak powyżej a symbolem d oznaczamy różnicę częstości empirycznej w dwóch grupach.

Przydatny link

(<http://biomath.info/power/index.htm>)

What is your study design?

Please indicate what type of procedure you need sample size / power calculations for:

one group

- ☐ paired t-test
- ☐ one-sample t-test
- ☐ chi-square test on proportion
- ☐ correlation

Go

two groups

- ☒ t-test on group means
- ☐ chi-square test on proportions

Tests for non-inferiority

- ☐ chi-square
- ☐ t-test

Go

Read about [Power Analysis](#)

Skorzystamy z gotowego programu (link podany na górze). Ten gotowy program pozwala nam ocenić liczebność niezbędną do odrzucenia hipotezy jeśli postawimy hipotezę dotyczącą jednej wartości oczekiwanej czy jednego prawdopodobieństwa lub dwóch wartości oczekiwanych czy dwóch prawdopodobieństw. To jest pierwszy wybór opcji. Jak widać wybrano weryfikację hipotezy o dwóch wartościach oczekiwanych.

Przydatny link
(<http://biomath.info/power/index.htm>)

Unpaired t-test

Find sample size:	Find effect size:
If you can estimate group means and standard deviation, use this form to find the number of subjects you need. Group 1 mean: 315.3 Group 2 mean: 283.4 Standard deviation: 219.3 Click here for sample size: Result You will need 744 subjects in Group 1 You will need 744 subjects in Group 2	If you know the number of subjects and the standard deviation of your measurement, use this form to see how small a difference you can detect. N for Group 1: Standard deviation: Click here for effect size: Result You can show a difference of size
For different power or significance level, change the fields below: Alpha: Prob(reject H_0 when H_0 is true) 0.05 Power: Prob(reject H_0 when H_1 is true) 0.80	
For unequal sample sizes, change the ratio below: Group2/Group1: 1.0	
Choose another study design v. 03/03 Contact: sh2@columbia.edu Source: G.W. Snedecor & W.G. Cochran Statistical Methods (7th ed., p. 104)	

Teraz czas na nasze założenia. Średnie w obydwu grupach, odchylenie standardowe wspólne dla wszystkich. Wybrane jest w okienku na dole jednakowa liczebność grup. Potem tylko poziom istotności i moc testu (tu 0,8 – słabszy test). I mamy podana wymagana liczebność. Wszystko jest funkcja różnicy poziomów grup – czyli jak bardzo czynnik różnicuje poziom cechy oraz odchylenia standardowego w grupach. W tym przykładzie powyżej założone jest bardzo duże odchylenie standardowe i dlatego takie duże liczebności prób. Zawsze znajdziemy taką liczebność która pozwoli nam odrzucić hipotezę, ale jeśli różnica między średnimi jest bardzo mała to liczebność będzie z reguły duża. Możemy czasem wpaść w logiczną pułkę i wykazywać przy pomocy weryfikacji hipotezy że różnica powiedzmy 2 kg w wydajności mlecznej różnych ras jest istotna, a wartość ta praktycznie mieści się w błędzie pomiaru.

Proszę spróbować samemu przewidywać liczebność dla sytuacji która jest Wam znana.

Przydatny link

(<http://biomath.info/power/index.htm>)

What is your study design?

Please indicate what type of procedure you need sample size / power calculations for:

one group

- ☐ paired t-test
- ☐ one-sample t-test
- ☐ chi-square test on proportion
- ☐ correlation

Go

two groups

- ☐ t-test on group means
- ☒ chi-square test on proportions

Tests for non-inferiority

- ☐ chi-square
- ☐ t-test

Go

Read about [Power Analysis](#)

Tu zastosowanie „liczydła” do porównania prawdopodobieństw. Zakłada się stosowanie testu Chi-kwadrat – słabszego od testu u zbudowanego na rozkładzie normalnym standaryzowanym.

Wybrane jest porównanie dwóch populacji, prawdopodobieństw w dwóch populacjach

Jak liczna musi być próba, aby wykazać różnice
w prawdopodobieństwie wystąpienia
pewnego patogenu w dwóch populacjach?
(0.3 / 0.4)

Find sample size:

**If you know the two proportions, use this form to find how many
subjects you need.**

Group 1 proportion: 0.3

Group 2 proportion: 0.4

Click here for sample size: [Result](#)

You will need 376 subjects in Group 1

You will need 376 subjects in Group 2

I założenia – częstości empiryczne w dwóch grupach oraz wynik czyli liczebność wymagana do odrz
Moje pytanie brzmi – czy to narzędzie szacuje liczebność prób przy założeniu jednostronnej czy dw

Podsumowanie

- Właściwe zaprojektowanie i zaplanowane doświadczenie daje duże szanse na weryfikację postawionej hipotezy w sposób wiarygodny przy wykorzystaniu minimalnej liczby zwierząt
- Wymaga to wiedzy na temat danego doświadczenia - na podstawie literatury lub badań pilotowych.

Podsumowując aby zaplanować doświadczenie trzeba naprawdę dużo wiedzieć. Jakiegokolwiek założenia niezgodne ze specyfiką cechy czy populacji spowodują że szacowana liczebność będzie obciążona błędem założenia. Zatem jeśli chcemy przewidywać liczebność najpierw musimy się dowiedzieć jak najwięcej o specyfice problemu, cechy, populacji itd..

Jedno co wiemy – warto planować doświadczenia w układach ortogonalnych i wato pamiętać o jednostronnych hipotezach alternatywnych (oczywiście tam gdzie są one możliwe).

