Tomasz Gliszczyński, St. Paul School in Warsaw

Czy dwuzłotówka jest symetryczna?

Część I

To jest projekt lekcji statystyki, na której przeprowadzone będzie **testowanie** pewnej konkretnej **hipotezy statystycznej**, czyli lekcja będzie
o planowaniu eksperymentu i o przygotowywaniu reguły decyzyjnej a priori (przed wykonaniem eksperymentu)

Czy macie przypadkiem w oczach taki obrazek. Rozpoczyna się mecz i sędzia rzuca monetą. Jak orzeł, to jedni, a jak reszka – to drudzy. Czy moneta jest naprawdę symetryczna?
Weźmy monetę dwuzłotową i „pstryknijmy" ją na gładkim stole tak, aby się zakręciła. Musi się tak mocno kręcić, aby wyglądała jak błyszcząca kulka. Kręci się i trochę wędruje. Potem kręci się coraz słabiej i pod koniec, jeszcze zanim całkiem upadnie, widać, czy do góry będzie orzeł, czy reszka. Spróbujmy zrobić to parę razy. Wyeliminujmy te przypadki, gdy moneta uderzy w przeszkodę lub gdy spadnie ze stołu. Przy odrobinie wprawy wystarczy mały stół, aby nie spadła. Stół powinien być gładki. Szkolne ławki wyłożone laminatem są do tego bardzo dobre.
Czy częściej będzie orzeł, czy reszka?
Zróbmy pierwszą turę eksperymentu. Może to wyglądać na przykład tak. Uczniowie na każdej ławce (a więc w parach) kręcą monetą dwuzłotową powiedzmy po przynajmniej pięć - dziesięć razy na każdej ławce.

> *Ta część lekcji jest potrzebna, aby, po pierwsze, uczniowie opanowali w miarę dobrze sztukę „kręcenia" monetą, a po drugie, aby mogli obejrzeć samo zjawisko, wstępnie je poznać i być może zainteresować się nim, stworzyć pierwszy wstępny, emocjonalny stosunek do problemu i postawić pierwsze hipotezy. Mogą to być bardzo różne hipotezy. Na przykład mogą być takie: gdy pstrykamy w orła, to wypadnie orzeł, a gdy pstrykamy w reszkę, to wypadnie reszka. Hipotezy mogą być najróżniejsze i z punktu widzenia dalszego planu lekcji mogą być bardzo niewygodne. Liczby poniżej dziesięciu mogą prowadzić do dość zaskakujących hipotez. Inną zaskakującą hipotezą, którą stawiali studjujący nauczyciele było na przykład to, że gdy orła nazwiemy sukcesem a reszkę porażką, to częsciej wypadnie orzeł i na odwrót – nazwanie reszki sukcesem powoduje, że częściej pojawiają się reszki. Muszę powiedzieć, że ten fragment lekcji bywa trudny. Masz zamiar pokazać* piękny kawałek matematyki, a tu wchodzisz w psychoparametanaukę. Niestety, czasem musisz być tutaj arbitralny i stanowczy(a) i przeforsować pierwszą hipotezę: **orzeł wypada częściej**. Możesz mieć oczywiście pecha. Generalnie jednak będziesz w większości. Będą stoliki, gdzie częściej wypadnie orzeł. To będą twoi sprzymierzeńcy. Będą stoliki, gdzie orzeł i reszka będą po równo. Ci będą najpewniej nieufni. Wiedzą skądinąd, że prawdopodobieństwo orła i reszki *są równe, a więc po pół. W końcu przecież nauczali tego sami przez czas jakiś, a ci, którzy ich nauczali, nauczali tego pewnie przez całe życie. No i będziesz miał(a) takich, którym orzeł wypadł rzadziej. Ci mogą wbrew pozorom najbardziej ci pomóc, bo oni najczęściej powiedzą, że trzeba przeprowadzić więcej niż 5 czy 10 eksperymentów.*

Stawiamy teraz hipotezę: **Orły wypadają częściej**.
Czy można tę hipotezę rozstrzygnąć przez eksperyment polegający na pięćdziesięciokrotnym powtórzeniu „pstryknięcia" monetą.

> *To ograniczenie do 50 ma dwie korzyści. Uczniowie będą pracowali parami przy ławkach, nie wiedząc, jakie wyniki uzyskają inne pary. A mogą to być przecież różne wyniki! Dopiero potem można tę różnorodność wyników spożytkować, zadają następne pytania. Jeszcze później można będzie je zsumować. Drugą korzyścią jest to, że do pięćdziesięciu powtórzeń w schemacie Bernoulliego nie trzeba poważniejszych urządzeń do liczenia silni niż kalkulator graficzny.*

Przyjęcie naszej hipotezy będzie jednocześnie odrzuceniem hipotezy, zwanej zerową, że moneta jest symetryczna.
Pewnym naprowadzeniem może być pytanie: co może nas zmusić do odrzucenia tej hipotezy zerowej, że moneta jest symetryczna?
Czy na przykład, wypadnięcie 50 orłów po kolei (zero reszek) spowoduje konieczność odrzucenia hipotezy zerowej na rzecz naszej hipotezy, zwanej hipotezą alternatywną, że prawdopodobieństwo wyrzucenia orła jest większe niż jedna druga? Przecież nawet wtedy, gdy prawdopodobieństwo orła byłoby tylko pół, to istnieje szansa na 50 orłów pod rząd. Co prawda to prawdopodobieństwo wynosi $0.5^{50} = 8.8817842 \times 10^{-16}$. Ryzyko, jakie można pominąć. No, ale gdyby wypadło tylko zero reszek lub jedna reszka, czyli conajwyżej jedna reszka, to czy dalej odrzucilibyśmy hipotezę o symetrii monety na rzecz hipotezy alternatywnej, że prawdopodobieństwo wypadnięcia orła jest większe niż pół? A co, gdy conajwyżej dwie?
I tak będziemy rozumować dalej...
Trzeba jeszcze zaplanować, jak praktycznie rejestrować wyniki. Do tego celu będzie służyć specjalnie zaprojektowany formularz, podany na str. 18. Jeden z uczniów pstryka monetą, a drugi rejestruje wynik, kreśląc odpo-

wiednio drogę od wierzchołka grafu, raz w lewo, raz w prawo, w zależności od tego jak upadła moneta.
Poniższe zestawienie pomoże podjąć decyzję.
Zakładamy hipotezę zerową (prawdopodobieństwo orła jest pół).

Prawdopodobieństwo	jest równe
50 orłów	$8.8817842 \times 10^{-16}$
49 lub 50 orłów	$4.4408921 \times 10^{-14}$
48, 49, lub 50 orłów	$1.13331566 \times 10^{-12}$
47, 48, 49 lub 50 orłów	$1.85416127 \times 10^{-11}$
46, 47, 48, 49 lub 50 orłów	$2.23089103 \times 10^{-10}$
45, 46, ... , 49 lub 50 orłów	$2.104926011 \times 10^{-9}$
44, 45, ..., 49 lub 50 orłów	$1.621870283 \times 10^{-8}$
43, 44, ..., 49 lub 50 orłów	$1.049338714 \times 10^{-7}$
42, 43, ..., 49 lub 50 orłów	$5.817779023 \times 10^{-7}$
41, 42, ..., 49 lub 50 orłów	$2.807050047 \times 10^{-6}$
40, 41, ..., 49 lub 50 orłów	$1.193066584 \times 10^{-5}$
39, 40, ..., 49 lub 50 orłów	$4.510745054 \times 10^{-5}$
38, 39, ..., 49 lub 50 orłów	$1.529320008 \times 10^{-4}$
37, 38, ..., 49 lub 50 orłów	$4.681114554 \times 10^{-4}$
36, 37, ..., 49 lub 50 orłów	.0013010857
35, 36, ..., 49 lub 50 orłów	.003300224
34, 35, ..., 49 lub 50 orłów	.0076733389
33, 34, ..., 49 lub 50 orłów	.0164195688
32, 33, ..., 49 lub 50 orłów	.0324543235
31, 32, ..., 49 lub 50 orłów	.0594602263
30, 31, ..., 49 lub 50 orłów	.1013193755
29, 30, ..., 49 lub 50 orłów	.1611181602
28, 29, ..., 49 lub 50 orłów	.2399438308
27, 28, ..., 49 lub 50 orłów	.3359055169
26, 27, ..., 49 lub 50 orłów	.4438624137
25, 26, ..., 49 lub 50 orłów	.5561375863
24, 25, ..., 49 lub 50 orłów	.6640944831
23, 24, ..., 49 lub 50 orłów	.7600561692
22, 23, ..., 49 lub 50 orłów	.8388818398
21, 22, ..., 49 lub 50 orłów	.8986806245
20, 21, ..., 49 lub 50 orłów	.9405397737
19, 20, ..., 49 lub 50 orłów	.9675456765
18, 19, ..., 49 lub 50 orłów	.9835804312
17, 18, ..., 49 lub 50 orłów	.9923266611
16, 17, ..., 49 lub 50 orłów	.996699776
15, 16, ..., 49 lub 50 orłów	.9986989143
14, 15, ..., 49 lub 50 orłów	.9995318885
13, 14, ..., 49 lub 50 orłów	.999847068
12, 13, ..., 49 lub 50 orłów	.9999548925
11, 12, ..., 49 lub 50 orłów	.9999880693
10, 11, ..., 49 lub 50 orłów	.999997193
9, 10, ..., 49 lub 50 orłów	.9999994182
8, 9, ..., 49 lub 50 orłów	.9999998951
7, 8, ..., 49 lub 50 orłów	.9999999838
6, 7, ..., 49 lub 50 orłów	.9999999979
5, 6, ..., 49 lub 50 orłów	.9999999998
4, 5, ..., 49 lub 50 orłów	$1-1.85416127 \times 10^{-11}$
3, 4, ..., 49 lub 50 orłów	$1-1.13331566 \times 10^{-12}$
2, 3, ..., 49 lub 50 orłów	$1-4.4408921 \times 10^{-14}$
1, 2, ..., 49 lub 50 orłów	$1-8.8817842 \times 10^{-16}$
0, 1, ..., 49 lub 50 orłów	1

Zanim przygotujemy się do podjęcia decyzji, a więc zanim ustalimy **regułę decyzyjną**, musimy określić ryzyko (prawdopodobieństwo) odrzucenia hipotezy zerowej, gdyby była ona prawdziwa. Gdybyśmy zdecydowali się na ryzyko co najwyżej $8.8817842 \times 10^{-16}$, to naszą regułą decyzyjną byłoby:
odrzucić hipotezę zerową, jeśli pojawią się same orły. Cóż, bardzo dobra reguła, tyle tylko, że mało praktyczna. Ta reguła decyzyjna raczej nie pozwoli nam na wykrycie tego, że prawdopodobieństwo orła jest większe niż pół. Bo nawet gdyby było ono 0.9, to prawdopodobieństwo przyjęcia decyzji, że orzeł jest bardziej prawdopodobny byłoby $0.9^{50} = 0.005$.
Musimy określić, na jakie ryzyko możemy sobie pozwolić. Jest to sprawa pozamatematyczna. O tym ryzyku decydować mogą najróżniejsze czynniki, na przykład **koszty złej decyzji**. Z drugiej strony są **koszty eksperymentu**. Tym razem postanowiliśmy zakręcić monetą pięćdziesiąt razy, ale może okazać się, że trzeba to zrobić więcej razy. Na razie jesteśmy związani z tą liczbą między innymi z powodów numeryczno-matematycznych (nie umiemy na razie liczyć zbyt dużych silni). Wracając do ryzyka. W statystyce to ryzyko nazywa się ***poziomem istotności*** i określa w procentach. Spróbujmy określić poziom istotności na 5% i dobierzmy do tego regułę decyzyjną. Spójrzmy na tabelkę. Widać, że kluczowy dla nas jest podkreślony wiersz. Jeśliby moneta była symetryczna, to prawdopodobieństwo wypadnięcia 32 lub 33 lub ....49 lub 50 orłów jest .0324543235, a więc mniej niż 5%, natomiast dołączenie 31 orłów spowoduje przekroczenioe 5% ryzyka. Tak więc nasza reguła decyzyjna jest taka:
Odrzucimy hipotezę zerową (że prawdopodobieństwo orła jest pół) na rzecz hipotezy alternatywnej (że prawdopodobieństwo orła jest większe niż pół), jeśli wypadnie 32 lub 33 lub ... 49 lub 50 orłów.
Ten zbiór {32, 33, 34, 35, 36, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50} nazywa się ***zbiorem odrzucenia hipotezy zerowej, a liczba 32 punktem krytycznym (przy poziomie istotności 5%).***
Cały ten proces tworzenia reguły decyzyjnej jest dla ludzi znających statystykę czymś bardzo rutynowym i ten dość długi proces zostałby streszczony na przykład tak:

H_0: p=0.5
[p jest prawdopodobieństwem wypadnięcia orła]
H_A: p>0.5
Poziom istotności 5%.
Planowany eksperyment polega na 50 "pstryknięciach" monetą.
Reguła decyzyjna:
Jeśli wypadnie 32 lub 33 lub ... 49 lub 50 orłów, to odrzucimy hipotezę zerową na rzecz alternatywnej.

Czy dobrze się rozumiemy? Jako ćwiczenie proszę spróbować wypełnić puste miejsca poniżej (zmienił się poziom istotności):

H_0: p=0.5
H_A: p>0.5
Poziom istotności 1%.
Planowany eksperyment polega na 50 "pstryknięciach" monetą.
Reguła decyzyjna:
Jeśli wypadnie___lub___lub ... 49 lub 50 orłów, to odrzucimy hipotezę zerową na rzecz alternatywnej.

Teraz możemy przeprowadzić eksperyment w klasie. Wygodnie jest skorzystać z arkusza z „drzewkiem”.
A oto jego wyniki w prawdziwej „klasie” złożonej z samych nauczycieli matematyki:
Wyniki (liczba sukcesów)

1. Elżbieta Gałązka	21
2. Teresa Rypina	22
3. Marek Grażyna	23
4. Grażyna Szacsznajder	26
5. Barbara Wierzchucka	26
6. Ewa Auguścik	26
7. Borzena Serzysko	26
8. Janina Ksionek	27
9. Rudzka Ewa	27
10. Mazek R.	27
11. Barbara Adamiak	27
12. Bogdan Maraszek	28
13. Jadwiga Klepacka	28
14. Maria Haska	28
15. Hanna Arciszewska	29
16. Abramowicz	29
17. B. Piasecka	29
18. Joanna Biłas	30
19. Alicja Zawadzka	31
20. Irena Głowa	32
21. Jolanta Zając	32
22. Elżbieta Kraska	32
23. Piotr Urbaniak	33
24. Grażyna Zalewska	33
25. Bożena Szczygielska	34
26. Iwona Chrząścik	34
27. Miłosława Wrzosek	35
28. Jadwiga Banasiuk	36
29. Anna Łaszczyk	36
30. Małgorzata Wolska	44

Ten ostatni wynik jest zaskakujący. Było w nim 29 orłów pod rząd! W naszej sytuacji nie był to jednak żaden argument, ponieważ **byliśmy w niewoli reguły decyzyjnej**. Teraz każda spośród 30 osób, niczego nie wiedząc o pozostałych, musi postąpić zgodnie z regułą decyzyjną. Tak więc 11 osób z numerami od 20 do 30 **odrzucają hipotezę zerową**. A co robi pierwsza dziewiętnastka? Każda z nich mówi to samo: **eksperyment nie pozwolił mi odrzucić hipotezy zerowej**, więc do pewnego stopnia **potwierdził ją**. Nie udowodnił jej coprawda, lecz również nie obalił.

Czy to jest w porządku, że **badając to samo zjawisko i przeprowadzając ten sam eksperyment eksperymentatorzy podejmują inne decyzje?** Czy to nie przeczy podstawowemu założeniu nauk przyrodniczych?

Rola statystyka jest rzeczywiście w świecie nauki wyjątkowa. To on planuje eksperyment i mówi, co należy zrobić i jak należy podjąć decyzję w przypadku takich to a takich wyników. **Eksperymentator**, mając już ten schemat postępowania, nie musi już interesować się matematycznym uzasadnieniem reguły decyzyjnej. On tylko **musi wykonać eksperyment, a potem zastosować się do wcześniej ustalonej reguły decyzyjnej**.

Przypuśćmy na chwilę, że moneta rzeczywiście jest symetryczna i że duża liczba ludzi zamówiła u statystyka ekspertyzę, jak przeprowadzić badanie. Poziom istotności 5% oznacza, że około 5% badaczy **odrzuci hipotezę**, że orzeł i reszka mają równe szanse, mimo że jest prawdziwa. Oznacza to że statystyk wie, że będzie miał kłopoty z około 5% klientów. (Przy eksperymencie z 50 powtórzeniami faktycznie z około 3.25% klientów. Patrz podkreślony wiersz w tabeli skumulowanych prawdopodobieństw.) Tak na prawdę, to statystyk ryzykuje jeszcze mniej. Najczęściej są jakieś dodatkowe argumenty na rzecz hipotezy alternatywnej. W naszym przypadku były to wstępne próby z monetą.

Wynik w naszej klasie daje dodatkowe argumenty, że coś z symetrią dwuzłotówki jest nie tak. Gdyby była symetryczna, to liczba osób odrzucających H_0 byłaby bliska 3.25%. U nas była to proporcja 11/30, czyli około 37%. **Czy dwuzłotówka rzeczywiście jest niesymetryczna?**

Doświadczenie na lekcji można wykonać i omówić w ciagu dwóch godzin lekcyjnych, ale potrzebna jest do tego **dobra organizacja pracy** i dobrze przygotowana rejestracja wyników. Załączam projekt formularza do takiej rejestracji. Na tym formularzu rejestruje się wyniki grubym kolorowym mazakiem, tak aby droga mazaka nakreślona zygzakiem pokazywała wyniki kolejnych zakręceń monetą. Najlepiej, jeżeli jedna osoba „kręci monetą” a druga rejestruje wyniki mazakiem, czyli praca wykonywana jest przez uczniów parami.

Pozostaje pytanie:
Jakie jest **prawdziwe prawdopodobieństwo**, że dwuzłotówka "pstryknieta" tak mocno, aby się zakręciła jak kuleczka, tak szybko, żeby nie było widać ani orła ani reszki, a po kilkunastu sekundach upadła orłem do góry?
O tym w następnej części.