

Zastosowanie testu statystycznego t-Studenta (*paired sample t-test*) w badaniu zachowań klientów sklepu e-commerce



Wojciech Moszczyński



Analiza danych może przynieść ogromne korzyści dla sklepów e-commerce. Podstawowym źródłem informacji o zachowaniu klientów w sklepie internetowym jest rejestr ich zakupów. Na tej podstawie możemy zaobserwować jak często klienci przychodzą do sklepu, które produkty wybierają, ile wydają, a nawet możemy dowiedzieć się jak reagują na zmiany cen. Analizowanie indywidualnych klientów w bazie, która zawiera na przykład 100 tys. rekordów jest mało efektywne. Dlatego tworzy się specjalne klastry, do których przyporządkowuje się klientów.

Przyporządkowanie klientów do klastrów

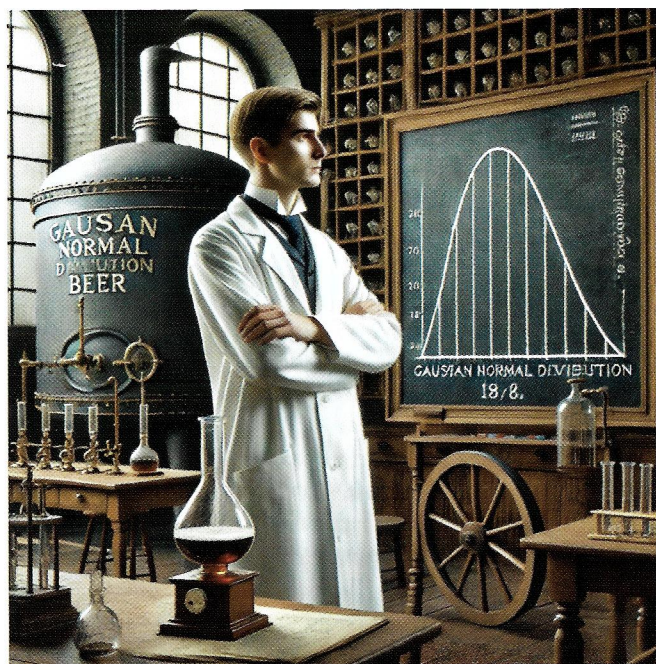
Wyobraźmy sobie taką sytuację, że mamy hurtownię techniczną e-commerce. Przychodzą do niej klienci i kupują produkty z różnych kategorii. Na podstawie rejestru sprzedaży możemy podzielić tych klientów na hydrauliczków, tynkarzy, elektryków oraz np. stolarzy i murarzy. Mając te pięć klastrów, możemy zacząć analizować zachowania klientów wewnątrz nich. Dzięki temu możemy zaobserwować, że murarze przychodzą częściej niż hydraulicy. Elektrycy są bardziej wrażliwi na zmiany cen, a tynkarze zostawiają w sklepie najwięcej pieniędzy.

To oczywiście bardzo uproszczony obraz analizy danych. Dzięki szczegółowym informacjom możemy stworzyć efektywną strategię sprzedaży oraz dobrze skonstruowaną dla każdego klastra promocję.

Nasuwa się pytanie, w jaki sposób porównywać zachowania tych pięciu specyficznych grup klientów? Analizę należy przeprowadzić za pomocą statystycznych testów t. Można wyróżnić trzy rodzaje tych testów t. Poświęcimy im najbliższe trzy publikacje.

Historia testu t-Studenta

Test t-Studenta, znany również jako test t, to jedno z najważniejszych narzędzi statystycznych stosowanych w analizie statystycznej danych. Jego twórcą jest William Sealy Gosset, student, który miał praktyki w browarze Guinness. Gosset musiał publikować swoje odkrycia pod pseudonimem „Student” z powodu ścisłych obostrzeń, jakie narzucał browar swoim pracownikom. W tamtym okresie, a był to początek XX wieku, przemysł browarniczy korzystał z wielu nowoczesnych technik, które dawały przewagę konkurencyjną firmom posiadającym odpowiednie know-how. Browar przywiązywał dużą wagę do



ochrony informacji dotyczących procesów produkcji, receptur, analiz jakościowych.

Aby spełnić warunki narzucone przez Guinnessa, Gosset zdecydował się na publikację swoich odkryć pod pseudonimem. W ten sposób zabezpieczył interesy firmy, ponieważ jego prawdziwe nazwisko nie było powiązane z publikacją, co uniemożliwiało identyfikację go jako pracownika Guinnessa. I tak ktoś kto mógłby mieć nazwisko równe Gaussa, Thomasa Bayesa czy Izaaka Newtona został w historii zapamiętany jako „student”; na cóż, tak czasem bywa.

Pozostał w cieniu, mimo że jego test, początkowo zauważony jedynie przez garstkę statystyków, z czasem zyskał uznanie i stał się jednym z najważniejszych narzędzi analizy danych na świecie.

Odkrycie Gosseta miało przełomowe znaczenie. Test *t* umożliwił analizy na małych próbkach, co było nieocenione nie tylko w przemyśle piwowarskim, ale także w medycynie, biologii, psychologii i wielu innych dziedzinach, gdzie często nie ma możliwości przeprowadzania badań na dużych zbiorach danych. Dzięki temu testowi można było wnioskować o różnicach między grupami nawet na podstawie ograniczonych danych.

Analiza klastra elektryków

Jak wspomniałem, analiza zachowań klientów na platformie sklepu e-commerce polega na analizie rejestru sprzedaży. Każda transakcja jest analizowana pod kątem celów jakie stawia sobie organizacja. Klientów nie analizuje się indywidualnie lecz grupuje się ich w klastry. W naszym przykładzie przyporządkowaliśmy klientów do pięciu grup zawodowych. Co ważne klientów można również analizować wewnątrz klastra przez tworzenie i porównywanie wewnętrznych wydzielonych grup.

Załóżmy, że mamy klastr elektryków. Wewnątrz tego klastra możemy rozróżnić np. klientów, którzy dużo kupują i tych, którzy kupują mało. Możemy ich podzielić na firmy z dużych i z małych miast i dalej szukać specjalnych cech w podgrupach.

Wyobraźmy sobie, że właściciel sklepu e-commerce postanowił zbadać jak reagują klienci z klastra elektrycy na zmniejszenie ilości promocji. Pracownicy sklepu zwrócili uwagę na to, że na promocję bardziej reagują mali klienci niż duże firmy. Zainteresowało to właściciela firmy, postanowił to zbadać. Podzielił firmę na małe, maksymalnie jedno- lub dwuosobowe działalności gospodarcze, oraz duże, które zatrudniają więcej niż dwie osoby i wydają dużo więcej pieniędzy. Proces jest łatwy do przeprowadzenia jeżeli analizuje się rejestr sprzedaży. Mając więc dwie grupy klientów wewnątrz klastra elektrycy właściciel postanowił dowiedzieć się, która grupa klientów będzie wrażliwa na spadek ilości promocji. Ewentualne wnioski z takiego badania mogą być kluczowe dla stworzenia efektywnej strategii sprzedaży. Jeżeli na przykład okazałoby się, że małe firmy są wrażliwe na ceny produktów należałoby amortyzować spadek ilości promocji poprzez przekazywanie małym firmom indywidualnych rabatów.

Generalnie właściciel firmy e-commerce zdecydował się już ograniczyć ilość promocji, ponieważ istotnie wpływała na kondycję jego przedsiębiorstwa. Jednocześnie chciał zapobiec odpływowi klientów z jego sklepu e-commerce.

Jak wybrać właściwy test *t*?

Zadanie polega na przeprowadzeniu analizy w celu sprawdzenia, w której grupie (małe czy duże firmy) ograniczenie promocji miało istotny wpływ na wielkość zakupów. Użyjemy niezależnego testu *t*-Studenta (*unpaired t-test*), aby porównać, czy spadek liczby zakupów po wprowadzeniu ograniczenia promocji jest istotny dla jednej z grup.

Test ten nie należy do trudnych; to zaledwie jedna linijka kodu. Jednak trzeba pamiętać o tym, aby wybrać właściwy test *t*.

Mamy trzy rodzaje testów *t*:

- One-sample *t*-test (test *t* dla jednej próby)
- Unpaired *t*-test (test *t* dla prób niezależnych)
- Paired sample *t*-test (test *t* dla prób zależnych).

One-sample *t*-test używamy, gdy chcemy porównać średnią z jednej próby z określoną wartością, np. średnią populacji. Np. sprawdzamy, czy średnia wysokość wydatków klastra elektryków różni się od ogólnej średniej wydatków wszystkich klientów sklepu.

Unpaired *t*-test używamy, gdy chcemy porównać średnie dwóch niezależnych grup, czyli takich, które nie mają powiązań między sobą. Porównujemy średnią wysokość wydatków klastra elektryków do klastra stolarzy. Możemy też porównać średnie wydatki elektryków kobiet i elektryków mężczyzn. Obie grupy są elektrykami, jednak są niezależne od siebie.

Paired sample *t*-test jest używany, gdy chcemy porównać średnie dla tych samych jednostek w dwóch różnych momentach lub warunkach. Porównujemy wyniki testów tej samej grupy elektryków przed i po szkoleniu, aby sprawdzić, czy szkolenie miało wpływ na stosowanie procedur bezpieczeństwa pracy.

W naszym przykładzie właściciel firmy chce sprawdzić, które firmy elektryczne (małe czy duże) są wrażliwe na ograniczenie ilości promocji. Małe i duże firmy to dwie oddzielne grupy klientów, które nie są ze sobą powiązane. Śledzimy zmiany

w poziomie zakupów tych samych firm przed i po promocji, porównujemy też ogólne różnice między dwoma grupami. Ponieważ analizujemy te same firmy przed i po promocjach musimy zastosować paired sample t-test.

Praktyczne wykorzystanie paired sample t-test

Paired sample t-test mierzy różnicę w wynikach w obrębie tej samej grupy w dwóch różnych momentach, dzięki czemu można uwzględnić, jak każda jednostka (firma) zmieniała swoje zachowanie pod wpływem wprowadzenia lub ograniczenia promocji. Dla każdej grupy klientów z osobna (małych firm oraz dużych firm) stawiamy następujące hipotezy:

Hipoteza zerowa (H0): Nie ma istotnej różnicy w wielkości zakupów przed i po wprowadzeniu ograniczenia promocji.

Hipoteza alternatywna (H1): Istnieje istotna różnica w wielkości zakupów przed i po wprowadzeniu ograniczenia promocji.

Paired sample t-test dla dwóch prób zależnych stosowany jest, gdy chcemy porównać dwie grupy zależnych danych (zakupów przed i po promocji) i sprawdzić, czy różnice w średnich są istotne statystycznie. Test ten zakłada, że dane są normalnie rozłożone, a wariancje są w przybliżeniu równe.

Ściągamy właściwe biblioteki w środowisku programowania Python.

```
[1]: import numpy as np
import scipy.stats as stats
```

Zakładamy, że mamy dane dotyczące liczby zakupów przez małe i duże firmy przed i po wprowadzeniu ograniczeń w promocjach.

```
[3]: # Symulacja danych dla małych i dużych firm: zakupy przed i po zmniejszeniu promocji
np.random.seed(42)

# Małe firmy: zakupy przed zmniejszeniem promocji (średnia = 22)
zakupy_przed_male_firmy = np.random.normal(loc=22, scale=5, size=100)

# Małe firmy: zakupy po zmniejszeniu promocji (średnia = 18)
zakupy_po_male_firmy = np.random.normal(loc=18, scale=5, size=100)

# Duże firmy: zakupy przed zmniejszeniem promocji (średnia = 27)
zakupy_przed_duze_firmy = np.random.normal(loc=27, scale=5, size=100)

# Duże firmy: zakupy po zmniejszeniu promocji (średnia = 26)
zakupy_po_duze_firmy = np.random.normal(loc=26, scale=5, size=100)
```

Ponieważ nie mamy danych historycznych, dla naszego przykładu generujemy losowe dane. Wygenerowaliśmy ilość zakupów przeprowadzonych przez 100 małych i 100 dużych firm.

Trzy pierwsze średnie zakupów tygodniowych dla firm dużych przed i po promocji wygląda następująco:

```
[4]: zakupy_przed_duze_firmy [:3]
[4]: array([28.7889368 , 29.80392263, 32.41525622])

[5]: zakupy_po_duze_firmy [:3]
[5]: array([21.85502495, 23.1990948 , 29.73646803])
```

Paired sample t-test dla firm – porównanie zakupów przed i po:

```
[6]: # Paired t-test dla małych firm (przed vs po promocji)
t_stat_male, p_value_male = stats.ttest_rel(zakupy_przed_male_firmy, zakupy_po_male_firmy)

# Paired t-test dla dużych firm (przed vs po promocji)
t_stat_duze, p_value_duze = stats.ttest_rel(zakupy_przed_duze_firmy, zakupy_po_duze_firmy)
```

Wydrukowanie wyników:

```
[7]: # Wyniki
print("Wyniki testu t dla małych firm:")
print(f"Statystyka t: {t_stat_male:.4f}, p-wartość: {p_value_male:.6f}")

print("\nWyniki testu t dla dużych firm:")
print(f"Statystyka t: {t_stat_duze:.4f}, p-wartość: {p_value_duze:.6f}")
```

Wyniki testu t dla małych firm:

Statystyka t: 4,8003, p-wartość: 0,000006

Wyniki testu t dla dużych firm:

Statystyka t: 1,1296, p-wartość: 0,261372

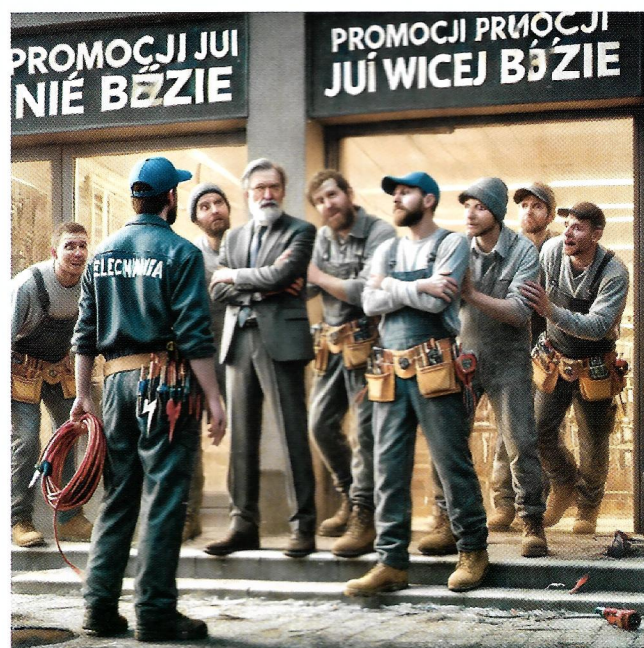
Paired t-test był przeprowadzany dla tych samych firm przed i po zmniejszeniu promocji, co pozwala na ocenę, czy zmiana jest statystycznie istotna. Jeśli p-wartość jest mniejsza od poziomu istotności ($\alpha = 0,05$), odrzucamy hipotezę zerową i uznajemy, że zmniejszenie promocji miało istotny wpływ na liczbę zakupów w tej grupie.

W naszym przykładzie małe firmy wykazują istotny spadek liczby zakupów po zmniejszeniu promocji, wiemy to ponieważ p-wartość $< 0,05$.

Duże firmy nie wykazują istotnej zmiany gdzie p-wartość $> 0,05$, co oznacza, że są niewrażliwe na zmniejszenie liczby promocji.

Wrażliwość na koniec częstych promocji

Jeżeli test t dla małych firm wykazuje istotną różnicę, oznacza to, że małe firmy są bardziej wrażliwe na brak promocji, co może oznaczać większe uzależnienie od rabatów. W takim



przypadku można rozważyć przywrócenie lub zwiększenie promocji specjalnie dla tej grupy. Duże firmy są mniej wrażliwe na promocje, a ich decyzje zakupowe są bardziej stabilne, niezależnie od zmian cenowych. Dla nich możemy zlikwidować promocje bez zagrożenia ich masowym odejściem.

Jeśli małe firmy reagują na brak promocji znacząco większym spadkiem zakupów, sugeruje to potrzebę rozważenia dla nich specjalnej strategii rabatowej, aby utrzymać ich lojalność i aktywność zakupową.

Podsumowanie

Test t nie cieszy się wielką popularnością wśród analityków danych. Wolał oni stosować często bardzo złożone modele analityczne, algorytmy shapleya¹ oraz złożone metody wyszukiwania korelacji. Jednak brak skutecznych i prostych metod statystycznych może okazać się zębne w skutkach.

Trzeba pamiętać o tym, że analiza zachowań klientów na platformach e-commerce składa się z tysięcy małych operacji, które jak mozaika tworzą prawdziwy obraz zachowań klientów. Gdyby właściciel platformy e-commerce zdecydował się na redukcję ilości rabatów, spotkałby się z masowym odpływem małych podmiotów w klastrze elektryków.

Dzięki prostemu testowi t dowiedział się jednak, że grupa małych firm jest wrażliwa na brak promocji, dzięki temu uchronił swój sklep przed stratami i utratą klientów.

Na podstawie wyników można wnioskować, które grupy klientów bardziej reagują na zmiany w promocjach, co pozwoliło na dostosowanie strategii marketingowej i cenowej do różnych wielkości firm, zwiększając efektywność działań promocyjnych. W ten sposób możemy lepiej zrozumieć, które grupy klientów w branży e-commerce są bardziej wrażliwe na zmiany w polityce promocji cenowych i jakie działania należy podjąć, aby minimalizować wpływ takich zmian na sprzedaż.

Wojciech Moszczyński

¹ <https://jakbadac dane.pl/co-to-jest-shapley-additive-explanation-s-shap/>

Wojciech Moszczyński – absolwent Katedry Ekonometrii i Statystyki Uniwersytetu Mikołaja Kopernika w Toruniu; specjalista z zakresu ekonometrii, finansów, data science oraz rachunkowości zarządczej. Specjalizuje się w optymalizacji procesów produkcyjnych i logistycznych. Prowadzi badania w obszarze rozwoju i zastosowania sztucznej inteligencji. Od lat zaangażowany w popularyzację machine learning oraz data science w środowiskach biznesowych.

„Śmietanka roślinna –
słodka i lekka
przyjemność!”



Delikatna, doskonale napowietrzona, zbliżona w smaku do tradycyjnej mlecznej śmietanki.

Z rozwiązaniami Hydrosol możesz stworzyć stabilny produkt, który bez kompromisu będzie rywalizował na rynku pod względem tekstury i smaku!

Nasza najnowsza receptura pozwala uzyskać:

- większą objętość
- większą stabilność
- przyjemniejszy smak
- zdecydowaną redukcję kosztów wytworzenia

hydrosol

We texture taste.

www.hydrosol.de