

Zastosowanie testu U Manna-Whitneya w analizie klientów branży e-commerce

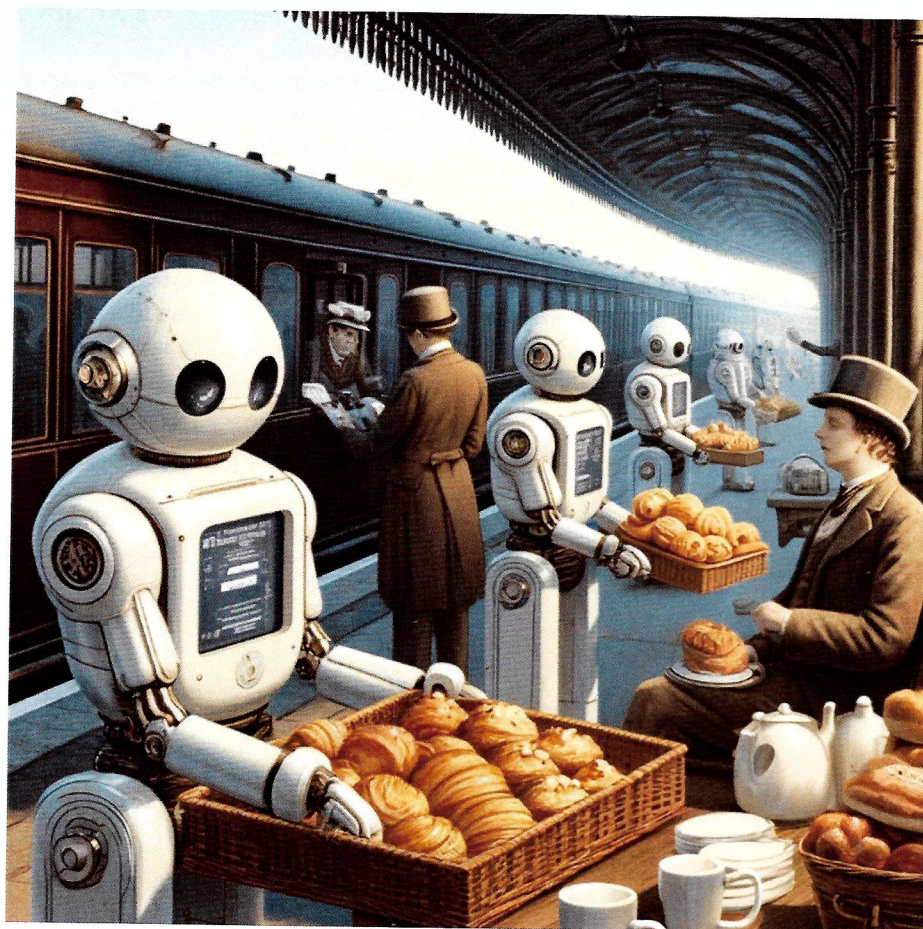
Test U Manna-Whitneya, znany również jako test Manna-Whitneya-Wilcoxon lub test rangowy Manna-Whitneya, został wynaleziony w 1947 roku przez dwóch statystyków – Henry'ego Manna i Donalda Whitneya. Test ten to nieparametryczna alternatywa dla testu t-Studenta. Test ten jest używany w sytuacjach, gdy dane nie spełniają założeń koniecznych do stosowania testów parametrycznych. Konkretnie chodzi o to, czy rozkład częstotliwości cech ma formę rozkładu normalnego. Innymi słowy jeżeli badana populacja jest statystycznie „dziwna”.

Co to znaczy, że rozkład cech spełnia założenia rozkładu normalnego?

W tym teście sprawdzenie zgodności populacji z rozkładem normalnym jest kluczowe. Gdy mówimy, że rozkład cech spełnia założenia rozkładu normalnego oznacza to, że wartości danej cechy (np. wydatki klientów, liczba zakupionych produktów, czas spędzony na stronie) są rozłożone w sposób charakterystyczny dla rozkładu normalnego. Taki rozkład ma specyficzne właściwości, które wpływają na sposób analizy danych. Założmy, że analizujemy kwoty wydawane przez klientów podczas zakupów. Jeśli te kwoty spełniają założenia rozkładu normalnego, to mają następujące cechy.

Symetria. Wydatki klientów są rozłożone symetrycznie wokół średniej wartości, czyli klienci wydają mniej więcej powyżej i poniżej średniej. Np., jeśli średnia kwota wydana przez klientów wynosi 100 zł, to tyle samo osób wydaje np. między 80 a 100 zł, co między 100 a 120 zł.

Jedna wartość środkowa. W rozkładzie normalnym średnia, mediana (wartość środkowa) i moda (najczęściej występująca wartość) są takie same. Dlatego ten rozkład jest symetryczny.



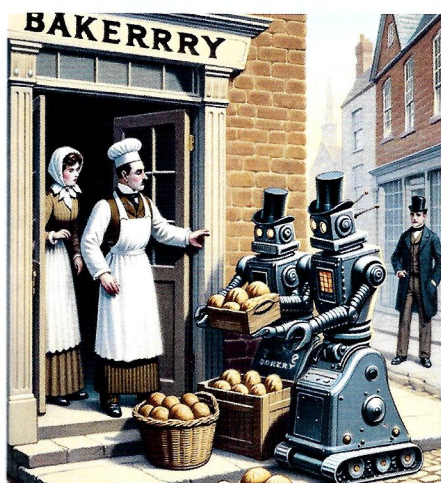
Gdy większość wydatków jest zbliżona do średniej to mamy wskazówkę, że wydatki klientów mają rozkład. Większość danych znajduje się blisko średniej kiedy około 68% wszystkich obserwowanych wydatków mieści się w przedziale jednego odchylenia standardowego od średniej, a 95% klientów w przedziale dwóch odchylen standardowych. Np., jeśli średnia kwota wydatków wynosi 100 zł, a odchylenie standardowe to 20 zł, to 68% klientów wydaje od 80 do 120 zł, a 95% od 60 do 140 zł.

Brak wartości odstających. W rozkładzie normalnym wartości ekstremalne (czyli bardzo małe lub bardzo duże wydatki) są rzadkie. Oznacza to, że większość klientów wydaje kwoty zbliżone do średniej, a przypadki, w których

ktoś wydaje bardzo mało lub bardzo dużo, są mało prawdopodobne.

Dlaczego rozkład normalny jest ważny?

Wiele standardowych testów statystycznych (np. test t-Studenta) zakłada, że dane mają rozkład normalny. Jeśli dane klientów sklepu internetowego (np. ich wydatki) są zgodne z rozkładem normalnym, można bezpiecznie stosować klasyczne testy parametryczne (czyli oparte na rozkładzie normalnym). Jeśli jednak dane są asymetryczne lub zawierają wiele wartości odstających, testy parametryczne mogą dawać błędne wyniki i w takich przypadkach trzeba stosować testy nieparametryczne, jak np. test U Manna-Whitneya.



Założmy, że chcemy sprawdzić, czy zmiana strategii rekomendacji produktów wpływa na średnią kwotę wydatków klientów. Jeśli rozkład wydatków klientów spełnia założenia normalności, możemy użyć testu t-Studenta, aby porównać średnie wydatki dwóch grup klientów (np. tych, którym polecano popularne produkty i tych, którym polecano produkty na podstawie historii zakupów). Zakładamy, że większość klientów wydaje kwoty blisko średniej, a ekstremalne wartości są rzadkie.

Jeśli jednak wydatki klientów są mocno zróżnicowane, np. niektórzy wydają bardzo mało, a inni bardzo dużo, co powoduje, że rozkład jest niesymetryczny, to założenia rozkładu normalnego nie są spełnione. W takim przypadku powinniśmy stosować testy nieparametryczne. Takim testem jest test U Manna-Whitneya.

Test U Manna-Whitneya

Test U Manna-Whitneya znajduje zastosowanie, gdy mamy do czynienia z danymi o nienormalnym rozkładzie lub gdy próbki danych mają różne rozmiary. Jest to często spotykane w analizach marketingowych, gdzie dane o zachowaniach klientów mogą być rozproszone, skrajne lub gdy dane pochodzą z niewielkich, nierównych grup.

W kontekście tworzenia systemów rekomendacji dla sklepów internetowych, test U Manna-Whitneya może być użyty do analizy różnic w zachowaniach dwóch grup klientów. Jest on przecież nieparametrycznym odpowiednikiem testu T, który operuje tylko na dwóch grupach. Wyobraźmy sobie sytuację, w której właściciel sklepu internetowego wprowadza

dwie różne strategii rekomendacji produktów dla klientów. Celem analizy jest ustalenie, która strategia generuje większe zyski, wyższe wskaźniki konwersji, bądź większe zaangażowanie użytkowników. Widzimy, że sposób rozumowania jest tutaj identyczny do podejścia jakie stosujemy w teście t-Studenta.

Założmy, że mamy dwie grupy klientów:

- Grupa A: klienci, którym rekomendacje wyświetlają się na podstawie popularności produktów.
- Grupa B: klienci, którym rekomendacje wyświetlają się na podstawie historii ich zakupów.

W takim przypadku można zastosować test U Manna-Whitneya, aby porównać medianę wskaźnika konwersji (procent klientów dokonujących zakupu po zobaczeniu rekomendacji) dla obu grup. Przykładem może być sytuacja, gdy grupa A ma bardziej rozproszony wskaźnik konwersji niż grupa B, co sugeruje, że niektóre rekomendacje są bardzo skuteczne, ale większość nie prowadzi do zakupów. Grupa B może mieć bardziej stabilne wyniki, co świadczy o bardziej spersonalizowanym podejściu.

Dlaczego test t-Studenta nie zawsze jest odpowiedni?

Test t-Studenta zakłada, że rozkład badanych zmiennych jest normalny, a wariancje w obu porównywanych grupach są równe. Jednak w praktyce, zwłaszcza w analizach marketingowych i badaniu wzorców zachowań klientów, te założenia często nie są spełnione. Przykładowo, wyniki zakupów czy kliknięć w systemach rekomendacji mogą być skrajnie różne w zależności od indywidualnych preferencji klientów, co prowadzi do dużej rozpiętości wyników. Takie dane mogą być asymetryczne lub zawierać wiele wartości odstających, co wyklucza zastosowanie testu t-Studenta.

Test U Manna-Whitneya jest odpowiedni w sytuacjach nierównych grup, rozproszenia wyników lub ich małej przewidywalności. To test nieparametryczny, więc nie wymaga spełnienia założeń dotyczących rozkładu normalnego. Test ten bazuje na rangach danych, a nie na ich wartości liczbowej, co sprawia, że jest bardziej odporny na odstępstwa od normalności oraz nierówność

wariancji. Właśnie dlatego często stosuje się go w analizach behawioralnych, w których dane pochodzą z naturalnych, często nieobliczalnych zachowań klientów, a rozkład wyników jest daleki od normalnego.

Analiza klientów internetowych za pomocą testu U

Mamy sklep internetowy, chcemy sprawdzić, czy nowa strategia marketingowa przyciąga bardziej zaangażowanych klientów. Firma podzieliła klientów na dwie grupy: jedni zobaczyli tradycyjną kampanię reklamową, a drudzy dostali spersonalizowane oferty bazujące na ich wcześniejszych zakupach. Aby ocenić skuteczność tych strategii, firma mierzy liczbę stron, które przeglądali klienci w obu grupach po wejściu na witrynę internetową sklepu.

Założmy, że w grupie tradycyjnej kampanii wyniki przeglądania stron są bardzo rozproszone – niektórzy klienci przeglądali bardzo wiele stron, inni zaś tylko jedną lub dwie. Z kolei w grupie, która otrzymała spersonalizowane oferty, większość klientów przeglądała stosunkowo małą liczbę stron, ale ich zachowanie jest bardziej jednolite. Mimo to zachowania te są z natury asymetryczne i rozproszone, co wyklucza zastosowanie testu t-Studenta.

Test U Manna-Whitneya pozwoliłby firmie porównać mediany liczby przeglądanych stron między grupami, uwzględniając zarówno nierówność wariancji, jak i potencjalne odstępstwa od rozkładu normalnego. Wynik tego testu mógłby wskazać, która strategia jest bardziej efektywna w przyciąganiu zainteresowania klientów.

Dlaczego test U Manna-Whitneya jest tak przydatny w analizie zachowań klientów?

Zachowania klientów często nie podlegają prostym wzorcom – są różnorodne, a dane na ich temat mogą być niestabilne i nieregularne. Dlatego testy nieparametryczne, takie jak test U Manna-Whitneya, są niezwykle przydatne w analizie rynku. Pomagają one lepiej zrozumieć, jak różne grupy klientów reagują na odmienne strategie marketingowe lub systemy rekomendacji, nawet jeśli dane są niejednolite lub mają dużą liczbę wartości odstających.