Bootcamp Data Science
Zajęcia 2

Przemysław Spurek

Testowanie hipotez

Testowanie hipotez

Obok **zagadnienia estymacji** drugim podstawowym działem wnioskowania statystycznego jest **wersyfikacja hipotez** (podejmowanie decyzji o prawdziwości lub fałszywości).

Testowanie hipotez

Obok **zagadnienia estymacji** drugim podstawowym działem wnioskowania statystycznego jest **wersyfikacja hipotez** (podejmowanie decyzji o prawdziwości lub fałszywości).

Definicja

Hipotezą statystyczną nazywamy każde przypuszczenie dotyczące nieznanego rozkładu badanej cechy populacji, o prawdziwości lub fałszywości, którego wnioskuje się na podstawie pobranej próbki.

Definicja

Statystyką nazywamy każdą zmienną losową będącą ustaloną funkcją próby losowej $X_1, X_2, \ldots, X_n$.

Testowanie hipotez

Definicja

Statystyką nazywamy każdą zmienną losową będącą ustaloną funkcją próby losowej $X_1, X_2, \ldots, X_n$.

Definicja

Najlepszą statystykę dla konkretnej hipotezy H nazywamy *statystyką testową* i oznaczamy $\delta(X_1, \ldots, X_n)$.

Testowanie hipotez

Definicja

Zbiór wartości statystyki testowej dzielimy na dopełniające się zbiory W oraz W' takie, że:

1. gdy $\delta(X_1, \ldots, X_n) \in W$ - hipotezę odrzucamy;
2. gdy $\delta(X_1, \ldots, X_n) \in W'$ - hipotezę przyjmujemy.

Testowanie hipotez

Definicja

Zbiór wartości statystyki testowej dzielimy na dopełniające się zbiory W oraz W' takie, że:

1. gdy $\delta(X_1, \ldots, X_n) \in W$ - hipotezę odrzucamy;
2. gdy $\delta(X_1, \ldots, X_n) \in W'$ - hipotezę przyjmujemy.

Zbiór W nazywamy zbiorem krytycznym lub zbiorem odrzuceń hipotez, a zbiór W' - zbiór przyjęć.

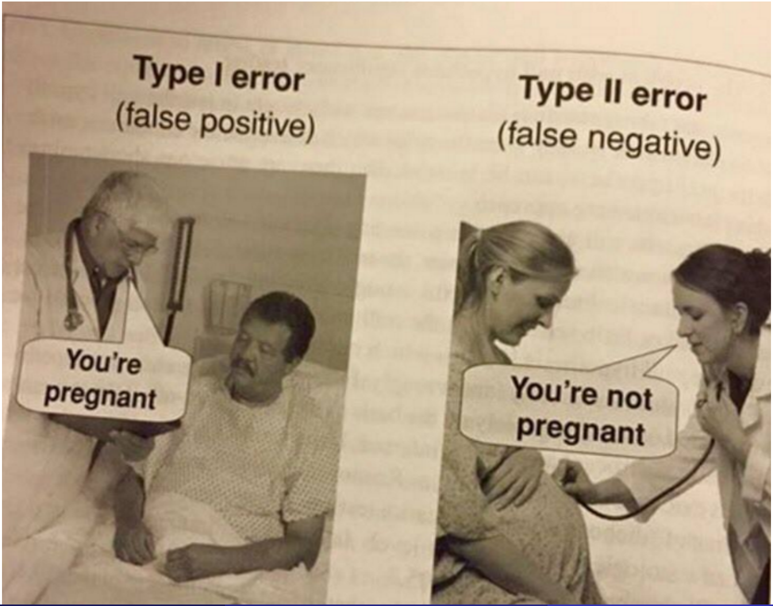
Testowanie hipotez

Weryfikując daną hipotezę H_0 za pomocą zaobserwowanej próbki możemy popełnić dwa podstawowe błędy

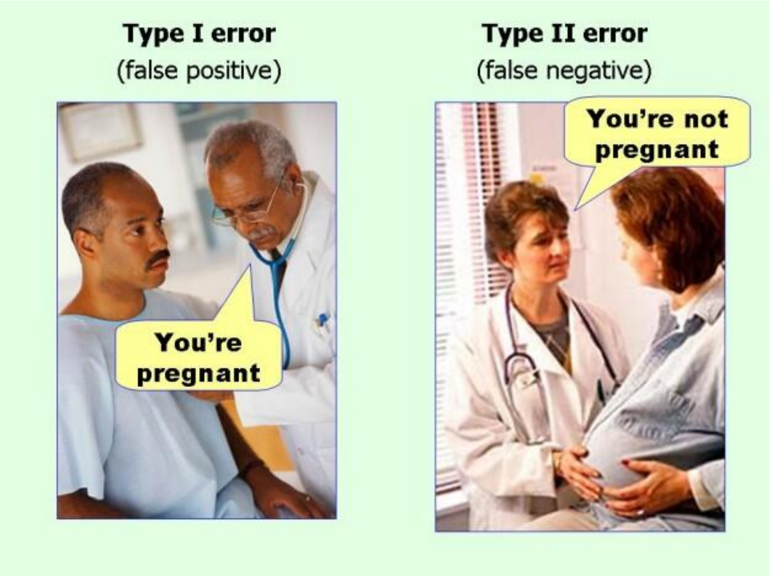
1. możemy odrzucić weryfikowaną hipotezę H_0 wtedy, gdy jest ona prawdziwa BŁĄD PIERWSZEGO RODZAJU;
2. możemy przyjąć weryfikowaną hipotezę H_0 jako prawdziwą, podczas gdy jest ona fałszywa BŁĄD DRUGIEGO RODZAJU.

Decyzja	Hipoteza H	
	jest prawdziwa	jest fałszywa
przyjąć weryfikowaną hipotezę H	decyzja poprawna	decyzja błędna (błąd drugiego rodzaju)
odrzucić hipotezę H	decyzja błędna (błąd pierwszego rodzaju)	decyzja poprawna

Testowanie hipotez



Testowanie hipotez



W testowaniu hipotez mogą wystąpić dwa rodzaje błędów

W kontroli jakości błąd typu I nazywany jest ryzykiem producenta, ponieważ odrzuca się produkt, mimo że spełnia wymagania regulacyjne.

W kontroli jakości błąd typu II nazywa się ryzykiem konsumenckim, ponieważ konsument otrzymuje element, który nie spełnia wymogów regulacyjnych.

W testowaniu hipotez mogą wystąpić dwa rodzaje błędów

Niektóre z bardziej mylących terminów w analizie statystycznej są **sensitivity** i **specificity**.
Tematy pokrewne to **positive predictive value** (PPV) i **negative predictive value** (NPV) testów statystycznych.

		Condition		
		Condition Positive	Condition Negative	
Test Outcome	Test Outcome Positive	**True Positive** (TP) = 25	**False Positive** (FP) = 175	**Positive predictive value=** = TP / (TP+FP) = 25 / (25+175) = 12.5%
	Test Outcome Negative	**False Negative** (FN) = 10	**True Negative** (TN) = 2000	**Negative predictive value=** = TN / (FN+TN) = 2000 / (10+2000) = 99.5%
		Sensitivity = = TP / (TP+FN) = 25 / (25+10) = 71%	**Specificitivity =** = TN / (FP+TN) = 2000 / (175+2000) = 92%	

W testowaniu hipotez mogą wystąpić dwa rodzaje błędów

Sensitivity Proporcja przykładów, które zostały poprawnie zidentyfikowane jako pozytywne przez test do wszystkich, które są pozytywne (powinny być poprawnie zidentyfikowane jako pozytywne).
Specificity Proporcja przykładów, które zostały poprawnie zidentyfikowane jako negatywne przez test do wszystkich, które są negatywne (powinny być poprawnie zidentyfikowane jako negatywne).
Positive Predictive Value (PPV) Proporcja pacjentów z pozytywnymi wynikami badania do tych, które zostały prawidłowo zdiagnozowane.
Negative Predictive Value (NPV) Proporcja pacjentów z ujemnymi wynikami badań do tych, które zostały negatywnie zdiagnozowane.

W testowaniu hipotez mogą wystąpić dwa rodzaje błędów

- Na przykład badania ciążowe mają wysoką sensitivity (czułość): gdy kobieta jest w ciąży, to prawdopodobieństwo pozytywnego wyniku badania jest bardzo wysokie.
- Natomiast wskaźnik ataku z użyciem broni atomowej na Warszawę powinien mieć bardzo dużą specificity: jeśli nie ma ataku to prawdopodobieństwo nieprzewidzenia tego jest małe.

Choć sensitivity i specificity charakteryzują test i są niezależne od częstości występowania, nie wskazują, która część pacjentów z nieprawidłowymi wynikami badania jest naprawdę nieprawidłowa. Te informacje są dostarczane przez PPV oraz NPV. Są to wartości istotne dla lekarza rozpoznającego pacjenta: kiedy pacjent ma pozytywny wynik badania, jak prawdopodobne jest, że pacjent jest w rzeczywistości chory?
Niestety wartości zależą od częstość występowania choroby.

Testowanie hipotez

Nie można zminimalizować obu rodzajów błędu. Więc ustalamy z góry poziom dopuszczalnego błędu pierwszego rodzaju – **poziom istotności**. Najczęściej ustalamy na $\alpha = 0.01$ albo $\alpha = 0.05$. Naszym celem jest zminimalizowanie błędu drugiego rodzaju.

Testowanie hipotez

Nie można zminimalizować obu rodzajów błędu. Więc ustalamy z góry poziom dopuszczalnego błędu pierwszego rodzaju – **poziom istotności**. Najczęściej ustalamy na $\alpha = 0.01$ albo $\alpha = 0.05$. Naszym celem jest zminimalizowanie błędu drugiego rodzaju.

W praktyce nie wyznacza się prawdopodobieństwa błędu drugiego rodzaju. Każdy test statystyczny (przy ustalonym poziomie istotności) posiada:

1. statystykę
2. zbiór krytyczny (zależy od poziomu istotności)

Testowanie hipotez

Jeżeli wartość statystyki testowej wpada do zbioru krytycznego, to odrzucamy hipotezę zerową na rzecz hipotezy alternatywnej.

Testowanie hipotez

Jeżeli wartość statystyki testowej wpada do zbioru krytycznego, to odrzucamy hipotezę zerową na rzecz hipotezy alternatywnej.

Jeżeli wartość statystyki testowej NIE wpada do zbioru krytycznego, to nie ma podstaw aby odrzucić hipotezę zerową.

Testowanie hipotez

Dla ustalonego poziomu istotności α - przy wykorzystaniu statystyki testowej $\delta(X_1, \ldots, X_n)$ - wyznaczamy taki zbiór krytyczny W, aby w przypadku, gdy hipoteza zerowa jest prawdziwa - spełniony był warunek

$$P(\delta(X_1, \ldots, X_n) \in W | H_0) = \alpha.$$

Testowanie hipotez

Dla ustalonego poziomu istotności α - przy wykorzystaniu statystyki testowej $\delta(X_1, \ldots, X_n)$ - wyznaczamy taki zbiór krytyczny W, aby w przypadku, gdy hipoteza zerowa jest prawdziwa - spełniony był warunek

$$P(\delta(X_1, \ldots, X_n) \in W | H_0) = \alpha.$$

Czasami osłabia się warunek (np. w przypadku zmiennych typu skokowego) do

$$P(\delta(X_1, \ldots, X_n) \in W | H_0) \leq \alpha.$$

Testowanie hipotez

Dla ustalonego poziomu istotności α - przy wykorzystaniu statystyki testowej $\delta(X_1, \ldots, X_n)$ - wyznaczamy taki zbiór krytyczny W, aby w przypadku, gdy hipoteza zerowa jest prawdziwa - spełniony był warunek

$$P(\delta(X_1, \ldots, X_n) \in W | H_0) = \alpha.$$

Czasami osłabia się warunek (np. w przypadku zmiennych typu skokowego) do

$$P(\delta(X_1, \ldots, X_n) \in W | H_0) \leq \alpha.$$

Można zauważyć, że ten zbiór krytyczny możemy wybrać na wiele sposobów (analogiczne rozumowanie jak w przypadku przedziałów największej wiarygodności).

Testowanie hipotez

Wobec wielu możliwości wyboru zbioru krytycznego najbardziej celowy byłby więc wybór zbioru W w taki sposób, aby minimalizował on prawdopodobieństwo błędu drugiego rodzaju.

Testowanie hipotez

Wobec wielu możliwości wyboru zbioru krytycznego najbardziej celowy byłby więc wybór zbioru W w taki sposób, aby minimalizował on prawdopodobieństwo błędu drugiego rodzaju.

Test, który przy ustalonym prawdopodobieństwie błędu pierwszego rodzaju minimalizuje prawdopodobieństwo błędu drugiego rodzaju nazywamy **testem najmocniejszym.**

Testowanie hipotez

W testach wykorzystywanych w praktyce bardzo często nie oblicza się błędu drugiego rodzaju, natomiast **przy założeniu prawdziwości weryfikowanej hipotezy** budujemy zbiór krytyczny w ten sposób, aby zagwarantować **małe prawdopodobieństwo** zaobserwowanie wartości **statystyki testowej należącej** do tego zbioru, równe z góry obranemu poziomowi istotności α.

Testowanie hipotez

W testach wykorzystywanych w praktyce bardzo często nie oblicza się błędu drugiego rodzaju, natomiast **przy założeniu prawdziwości weryfikowanej hipotezy** budujemy zbiór krytyczny w ten sposób, aby zagwarantować **małe prawdopodobieństwo** zaobserwowanie wartości **statystyki testowej należącej** do tego zbioru, równe z góry obranemu poziomowi istotności α.

Jeżeli wartość statystyki testowej wpadnie do przedziału krytycznego, to możemy stwierdzić, że zaszło zdarzenie o małym prawdopodobieństwie i wówczas **hipotezę należy odrzucić**.

Testowanie hipotez

W testach wykorzystywanych w praktyce bardzo często nie oblicza się błędu drugiego rodzaju, natomiast **przy założeniu prawdziwości weryfikowanej hipotezy** budujemy zbiór krytyczny w ten sposób, aby zagwarantować **małe prawdopodobieństwo** zaobserwowanie wartości **statystyki testowej należącej** do tego zbioru, równe z góry obranemu poziomowi istotności α.

Jeżeli wartość statystyki testowej wpadnie do przedziału krytycznego, to możemy stwierdzić, że zaszło zdarzenie o małym prawdopodobieństwie i wówczas **hipotezę należy odrzucić**.

Jeżeli wartość statystyki nie znajduje się w zbiorze krytycznym, to możemy jedynie twierdzić, że **nie ma podstaw do odrzucenia weryfikowanej hipotezy**

Testowanie hipotez

Testy hipotez statystycznych można podzielić na dwie grupy:

- testy parametryczne,
- testy nieparametryczne.

Testowanie hipotez

Testy hipotez statystycznych można podzielić na dwie grupy:

- testy parametryczne,
- testy nieparametryczne.

Testy parametryczne zakładają, że dane można dobrze opisać rozkładem, który jest określony przez jeden lub więcej parametrów, w większości przypadków przez rozkład normalny. Dla danego zestawu danych określane są parametry dopasowane do tej dystrybucji wraz z ich przedziałami ufności.

Testowanie hipotez

Testy hipotez statystycznych można podzielić na dwie grupy:

- testy parametryczne,
- testy nieparametryczne.

Testy parametryczne zakładają, że dane można dobrze opisać rozkładem, który jest określony przez jeden lub więcej parametrów, w większości przypadków przez rozkład normalny. Dla danego zestawu danych określane są parametry dopasowane do tej dystrybucji wraz z ich przedziałami ufności.

To podejście działa tylko wtedy, gdy dany zestaw danych jest właściwie przybliżony przez wybrany rozkład prawdopodobieństwa. Jeśli nie, wyniki testu parametrycznego mogą być całkowicie błędne. W takim przypadku należy użyć testów nieparametrycznych, które są mniej wrażliwe na dopasowanie rozkładu.

p-value

P-wartość, wartość p, prawdopodobieństwo testowe (ang. p-value, probability value) – prawdopodobieństwo, że zjawisko jakie zaobserwowano w jakimś pomiarze na losowej próbie statystycznej z populacji, mogło wystąpić przypadkowo, wskutek losowej zmienności prób, w sytuacji w której w populacji takie zjawisko wcale nie występuje.

p-value

P-wartość, wartość p, prawdopodobieństwo testowe (ang. p-value, probability value) – prawdopodobieństwo, że zjawisko jakie zaobserwowano w jakimś pomiarze na losowej próbie statystycznej z populacji, mogło wystąpić przypadkowo, wskutek losowej zmienności prób, w sytuacji w której w populacji takie zjawisko wcale nie występuje.

p-value - jest definiowane ściśle jako prawdopodobieństwo kumulatywne wylosowania próby **takiej, lub bardziej skrajnej, jak zaobserwowana**, przy założeniu, że **hipoteza zerowa jest spełniona.**

p-value

- Innymi słowy p-value określa, jak prawdopodobne jest uzyskanie danej wartości jako ekstremalnej przy założeniu, że hipoteza zerowa jest prawdziwa.
- Wartość, względem której porównywana jest p-value nazywamy poziomem istotności, najczęściej ustalany na poziomie 0.05.
- Taki sposób postępowania w celu sprawdzenia hipotezy nazywa się **wnioskowaniem statystycznym**.
- Jeśli
  - p-value jest mniejsza niż $p < 0.05$, to odrzucamy hipotezę zerową,
  - p-value jest większe niż $p > 0.05$, to nie mamy podstaw do odrzucenia hipotezy zerowej na zadanym poziomie istotności.

Ogólna procedura testowania hipotez

Ogólna procedura testowania hipotez:

- Musimy pobrać próbkę z danej populacji (W naszym przykładzie losowa próbka jest zdana).
- Musimy sformułować hipotezę zerową.
- Musimy obliczyć statystykę testową, której znamy rozkład prawdopodobieństwa.
- Musimy wykonać jedno z poniższych porównań:
  - zaobserwowaną wartość statystyki i przyjęty przedział krytyczny,
  - p-value i poziom istotności α.

Testowanie hipotez dotyczących wartości średniej

Model 1.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **znanym σ**.
Weryfikujemy hipotezę:

- H_0: $\mu = \mu_0$
- H_1: $\mu = \mu_1 > \mu_0$

W tym teście używa się statystyki:

$$U = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$[u(1-\alpha), +\infty)$$

gdzie $u(\alpha)$ jest kwantylem rozkładu normalnego.

Testowanie hipotez dotyczących wartości średniej

Model 2.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$, przy **znanym σ**.
Weryfikujemy hipotezę:

- H_0: $\mu = \mu_0$
- H_1: $\mu = \mu_1 < \mu_0$

W tym teście używa się statystyki:

$$U = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$(-\infty, -u(\alpha)]$$

gdzie $u(\alpha)$ jest kwantylem rozkładu normalnego.

Model 3
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **znanym** σ.
Weryfikujemy hipotezę:

- H_0: $\mu = \mu_0$
- H_1: $\mu = \mu_1 \neq \mu_0$

W tym teście używa się statystyki:

$$U = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$\left(-\infty, -u(1 - \frac{1}{2}\alpha)\right] \cup \left[u(1 - \frac{1}{2}\alpha), +\infty\right)$$

gdzie $u(\alpha)$ jest kwantylem rozkładu normalnego.

Testowanie hipotez dotyczących wartości średniej

https://github.com/przem85/bootcamp/blob/master/statistics/D06_Z01.ipynb

Zadanie

Z populacji, w której badana cecha ma rozkład $N(\mu, 4)$ wylosowano próbkę złożoną z 9 obserwacji. Na poziomie istotności $\alpha = 0.05$ zweryfikować hipotezę

1. H_0: $\mu = 2$ przy hipotezie alternatywnej H_1: $\mu < 2$,
2. H_0: $\mu = 2$ przy hipotezie alternatywnej H_1: $\mu > 2$,
3. H_0: $\mu = 2$ przy hipotezie alternatywnej H_1: $\mu \neq 2$,

jeżeli średnia z punktów wynosi $\bar{X} = 1.4$.

Testowanie hipotez dotyczących wartości średniej

Model 4.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **obu parametrach nieznanych**.
Weryfikujemy hipotezę:

- H_0: $\mu = \mu_0$,
- H_1: $\mu = \mu_1 > \mu_0$.

W tym teście używa się statystyki

$$t = \frac{\bar{X} - \mu_0}{S}\sqrt{n-1}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$[t(1-\alpha, n-1), +\infty)$$

gdzie $t(\alpha, n)$ jest kwantylem rozkładu t-Studenta przy n stopniach swobody oraz $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Model 5.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **obu parametrach nieznanych**.
Weryfikujemy hipotezę:

- H_0: $\mu = \mu_0$,
- H_1: $\mu = \mu_1 < \mu_0$.

W tym teście używa się statystyki:

$$t = \frac{\bar{X} - \mu_0}{S}\sqrt{n-1}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$(-\infty, -t(1-\alpha, n-1)]$$

gdzie $t(\alpha, n)$ jest kwantylem rozkładu t-Studenta przy n stopniach swobody oraz $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Model 6.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **obu parametrach nieznanych**.
Weryfikujemy hipotezę:

- H_0: $\mu = \mu_0$,
- H_1: $\mu = \mu_1 \neq \mu_0$.

W tym teście używa się statystyki

$$t = \frac{\bar{X} - \mu_0}{S}\sqrt{n-1}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$(-\infty, -t(1 - \frac{1}{2}\alpha, n-1)] \cup [t(1 - \frac{1}{2}\alpha, n-1), +\infty)$$

gdzie $t(\alpha, n)$ jest kwantylem rozkładu t-Studenta przy n stopniach swobody oraz $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Testowanie hipotez dotyczących wartości średniej

https://github.com/przem85/bootcamp/blob/master/statistics/D06_Z02.ipynb

Zadanie

W celu ustalenia, czy dotychczasowa norma okresu użytkowania ubrań ochronnych – wynosząca 150 dni – nie jest zbyt wysoka, zbadano faktyczny okres użytkowania ich na przykładzie 65 losowo wybranych robotników pracujących w normalnych warunkach. Otrzymano średnią długość okresu użytkowania 139 dni oraz odchylenie standardowe (S) 9.8 dni. Zakładając, że czas użytkowania ubrań ma rozkład normalny, stwierdzić, na poziomie istotności $\alpha = 0.01$, czy uzyskane wyniki stanowią podstawę do:

1. zmiany normy,
2. zmniejszenia normy,
3. zwiększenia normy.

One Sample t-Test for a Mean Value

- Aby zweryfikować hipotezę odnoszącą się do średniej dla danych pochodzących z rozkładu normalnego względem wartości referencyjnej, zwykle używamy **One Sample t-Test**, który jest oparty na rozkładzie t-Student.
- Jeśli znamy średnią i odchylenie standardowe, możemy obliczyć odpowiadający mu standardowy błąd i użyć wartości z rozkładu normalnego, aby ustalić prawdopodobieństwo znalezienia określonej wartości.
- W praktyce **One Sample t-Test** jest jednym z najczęściej spotykanych testów.

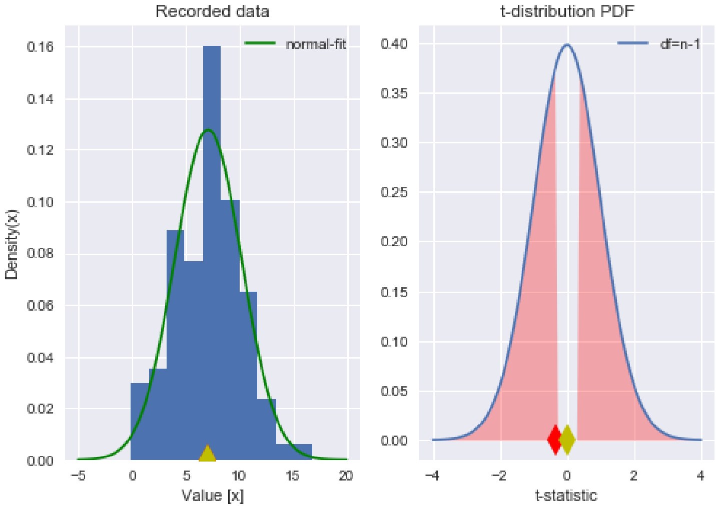
One Sample t-Test for a Mean Value

https://github.com/przem85/bootcamp/blob/master/statistics/D06_Z03.ipynb

Przykład

- Wygenerujmy 100 elementową próbkę z rozkładu normalnego ze średnią 7 i odchyleniem standardowym 3.
- Średnia z próbki jest bliska ale różna od rzeczywistej średniej.
- Zapomnijmy, o tym, że wiemy z jakiego rozkładu pochodzi próbka i wykonajmy **One Sample t-Test** dwoma sposobami:
  - na piechotę,
  - wykonując funkcję `stats.ttest_1samp`.

One Sample t-Test for a Mean Value



Weryfikacja hipotezy dotyczącej wariancji

Model 1.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **obu parametrach nieznanych** ($n < 50$).
Weryfikujemy hipotezę:

- H_0: $\sigma^2 = \sigma_0^2$,
- H_1: $\sigma = \sigma_1 > \sigma_0$.

W tym teście używa się statystyki:

$$\chi^2 = \frac{nS^2}{\sigma_0^2}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$[\chi^2(1-\alpha, n-1), +\infty)$$

gdzie $\chi^2(\alpha, n)$ jest kwantylem rozkładu χ^2 przy n stopniach swobody oraz $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Weryfikacja hipotezy dotyczącej wariancji

Model 2.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **obu parametrach nieznanych** ($n < 50$).
Weryfikujemy hipotezę:

- H_0: $\sigma = \sigma_0$,
- H_1: $\sigma = \sigma_1 < \sigma_0$.

W tym teście używa się statystyki:

$$\chi^2 = \frac{nS^2}{\sigma_0^2}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$(0, \chi^2(\alpha, n-1)]$$

gdzie $\chi^2(\alpha, n)$ jest kwantylem rozkładu χ^2 przy n stopniach swobody oraz $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Weryfikacja hipotezy dotyczącej wariancji

Model 3.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **obu parametrach nieznanych** ($n < 50$).
Weryfikujemy hipotezę:

- H_0: $\sigma = \sigma_0$,
- H_1: $\sigma = \sigma_1 \neq \sigma_0$.

W tym teście używa się statystyki:

$$\chi^2 = \frac{nS^2}{\sigma_0^2}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$(0, \chi^2(\frac{1}{2}\alpha, n-1)] \cup [\chi^2(1 - \frac{1}{2}\alpha, n-1), +\infty)$$

gdzie $\chi^2(\alpha, n)$ jest kwantylem rozkładu χ^2 przy n stopniach swobody oraz $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Weryfikacja hipotezy dotyczącej wariancji

https://github.com/przem85/bootcamp/blob/master/statistics/D06_Z04.ipynb

Zadanie

W celu oszacowania dokładności pomiarów wykonanych pewnym przyrządem dokonano 8 pomiarów pewnej wielkości i otrzymano: $18.17, 18.21, 18.05, 18.14, 18.19, 18.22, 18.06, 18.08$.
Zweryfikować na poziomie istotności $\alpha = 0.05$ hipotezę $\sigma^2 = 0.06$ wobec hipotez alternatywnych:

1. $\sigma^2 \neq 0.06$,
2. $\sigma^2 < 0.06$,
3. $\sigma^2 > 0.06$.

Weryfikacja hipotezy dotyczącej wariancji

Model 4.
Badana cecha X populacji generalnej ma rozkład $N(\mu, \sigma)$ przy **obu parametrach nieznanych** ($n \geq 50$).
Weryfikujemy hipotezę:

- H_0: $\sigma = \sigma_0$
- H_1: $\sigma = \sigma_1 > \sigma_0$

W tym teście używa się statystyki:

$$U = \sqrt{\frac{2nS^2}{\sigma_0^2}} - \sqrt{2n-3}$$

Zbiór krytyczny jest dany za pomocą przedziału:

$$[u(1-\alpha), +\infty)$$

gdzie $u(\alpha)$ jest kwantylem rozkładu normalnego.

Weryfikacja hipotezy dotyczącej wariancji

Analogicznie dla innych alternatywnych hipotez:
Model 5.
Model 6.

Weryfikacja hipotezy dotyczącej wariancji

https://github.com/przem85/bootcamp/blob/master/statistics/D06_Z05.ipynb

Zadanie

Do tarczy oddano 50 strzałów. Mierząc odległości trafień od środka tarczy, okazało się, że wariancja tych odległości jest równa $S^2 = 107,3\ cm^2$. Zakładając, że te odległości mają rozkład normalny na poziomie istotności $\alpha = 0.05$, zweryfikować hipotezę H, że wariancja odległości trafienia od środka tarczy jest równa $\sigma^2 = 100\ cm^2$ przy hipotezie alternatywnej:

1. $\sigma^2 \neq 100$,
2. $\sigma^2 > 100$,
3. $\sigma^2 < 100$.

Two Sample t-Test for a Mean Value

Test ten weryfikuje równość średnich dla dwóch zmiennych. Można to zapisać za pomocą zestawu hipotez:

- H_0: $\mu_1 = \mu_2$,
- H_1: $\mu_1 \neq \mu_2$,

gdzie μ_1, μ_2, to odpowiednio średnie z próby dla pierwszej i drugiej zmiennej.

Two Sample t-Test for a Mean Value

https://github.com/przem85/bootcamp/blob/master/statistics/D06_Z06.ipynb

Z uwagi na możliwość zdefiniowania grup, których owe zmienne się tyczą, wyróżnia się dwa przypadki:

- test na równość średnich dla grup niezależnych,
- test na równość średnich dla grup zależnych (powiązanych).

Zastosowanie testu dwóch średnich wymaga spełnienia przez zmienne warunku normalności rozkładu (w każdej analizowanej podgrupie).

Wszystkie poznane dotychczas testy zakładały, że dane pochodzą z rozkładu normalnego.

Wszystkie poznane dotychczas testy zakładały, że dane pochodzą z rozkładu normalnego.

Jak sprawdzić, że tak jest?
Co wtedy, gdy tak nie jest?