Statistical Tests

Tuesday 29th November, 2016

Statistical Tests

- up to this point
- basic procedure
- in practice - specific tests used

One and Two Tailed Tests

Experiment I

- test hypothesis - crime rates are $\boxed{\text{elevated}}$ when Internet services are down
- what are the appropriate alternative and null hypotheses ?

One and Two Tailed Tests

Israeltech challenge

Experiment I

- test hypothesis - crime rates are $\boxed{\text{elevated}}$ when Internet services are down
- what are the appropriate alternative and null hypotheses ?
- C_r crime rate
- C_r^{no-net} crime rate when Internet is down

$$H_0 : C_r = C_r^{no-net}$$
$$H_1 : C_r < C_r^{no-net}$$

One and Two Tailed Tests

Experiment II

- test hypothesis - crime rates are $\boxed{\text{changed}}$ when Internet services are down
- what are the appropriate alternative and null hypotheses ?

One and Two Tailed Tests

IsraelTech challenge

Experiment II

- test hypothesis - crime rates are $\boxed{\text{changed}}$ when Internet services are down
- what are the appropriate alternative and null hypotheses ?
- C_r crime rate
- C_r^{no-net} crime rate when Internet is down

$$H_0 : C_r = C_r^{no-net}$$
$$H_1 : C_r \neq C_r^{no-net}$$

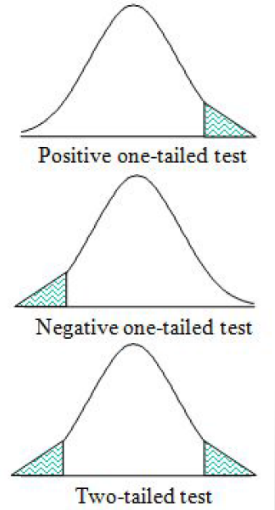
One and Two Tailed Tests

One tailed test

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta > \theta_0$$

Two tailed test

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$



board

The Two Tailed Coin inspector

test if a coin is fair with significance $\alpha = 0.1$:

- toss coin 10 times - 8 heads
- do you reject?
- what is the P-Value?

The Two Tailed Coin inspector

if the test is two tailed

$$H = \text{number of heads}$$

$$Pval = 2 \cdot \min\left\{P\left(H \geq 8\right), P\left(H \leq 8\right)\right\}$$

The Two Tailed Coin inspector

if the test is two tailed

$$H = \text{number of heads}$$

$$\begin{aligned} Pval &= 2 \cdot \min\left\{P\left(H \geq 8\right), P\left(H \leq 8\right)\right\} \\ &= 2 \cdot P\left(H \geq 8\right) \end{aligned}$$

The Two Tailed Coin inspector

Israeltech challenge

if the test is two tailed

$$H = \text{number of heads}$$

$$\begin{aligned} Pval &= 2 \cdot \min \left\{ P\left(H \geq 8\right), P\left(H \leq 8\right) \right\} \\ &= 2 \cdot P\left(H \geq 8\right) \\ &= 2 \cdot \left(\binom{10}{2} + \binom{10}{1} + 1 \right) \cdot 0.5^{10} \end{aligned}$$

The Two Tailed Coin inspector

if the test is two tailed

$$H = \text{number of heads}$$

$$\begin{aligned} Pval &= 2 \cdot \min \left\{ P\left(H \geq 8\right), P\left(H \leq 8\right) \right\} \\ &= 2 \cdot P\left(H \geq 8\right) \\ &= 2 \cdot \left(\binom{10}{2} + \binom{10}{1} + 1 \right) \cdot 0.5^{10} \\ &= 2 \cdot 0.055 \\ &= 0.109 > 0.1 \end{aligned}$$

Rejection Region

Automation

two tailed coin inspector

- $H = 8$ - $Pval = 0.109$ - **not rejected**
- $H = 9$ - $Pval = 0.021$ - **rejected**

instead of computing P-Value for each coin we can safely reject...

Rejection Region

Automation
two tailed coin inspector

- $H = 8$ - $Pval = 0.109$ - **not rejected**
- $H = 9$ - $Pval = 0.021$ - **rejected**

instead of computing P-Value for each coin we can safely reject...

$$R = \{X : or\,(H_X \geq 9, H_X \leq 1)\}$$
$$= \{X : |5 - H_X| \geq 4\}$$

Rejection Region

$$\boxed{R = \{X : T(X) > c\}}$$

- T - test statistic
- c - critical value
- $P(T(X) > c) = ?$

Rejection Region

$$\boxed{R = \{X : T(X) > c\}}$$

- T - test statistic
- c - critical value
- $P(T(X) > c) = \alpha$

One Tailed Test Example

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is known
- $H_0 : \mu \leq 0$
- $H_1 : \mu > 0$

One Tailed Test Example

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is known
- $H_0 : \mu \leq 0$
- $H_1 : \mu > 0$
- **Test**: $R = \{X : T\left(X\right) > c\}$
- $T\left(X\right) = \frac{1}{n}\sum_{i=1}^{n} X_i$

One Tailed Test Example

$$T\left(X\right)\sim\mathcal{N}\left(0,\frac{\sigma^{2}}{n}\right)$$

why?

One Tailed Test Example

Israeltech challenge

$$T\left(X\right) \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right)$$

why?

- $X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$, $Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$
- X and Y are independent
- $\alpha \in \mathbb{R}$

One Tailed Test Example

$$T(X) \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right)$$

why?

- $X \sim \mathcal{N}\left(\mu_X, \sigma_X^2\right)$, $Y \sim \mathcal{N}\left(\mu_Y, \sigma_Y^2\right)$
- X and Y are independent
- $\alpha \in \mathbb{R}$

$$X + Y \sim \mathcal{N}\left(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2\right)$$
$$\alpha \cdot X \sim \mathcal{N}\left(\alpha\mu_X, \alpha^2\sigma_X^2\right)$$

One Tailed Test Example

Israeltech challenge

$$T(X) \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right)$$

why?

$$T(X) = \frac{1}{n}\left[X_1 + X_2 + \ldots + X_n\right]$$
$$T(X) \sim \mathcal{N}\left(0, \frac{1}{n^2} n \sigma^2\right)$$

One Tailed Test Example

- Z is standard normal - $Z \sim \mathcal{N}(0,1)$
- Φ the CDF of Z

$$P(T > c) = P\left(\frac{T-\mu}{\sigma/\sqrt{n}} > \frac{c-\mu}{\sigma/\sqrt{n}}\right)$$

One Tailed Test Example

- Z is standard normal - $Z \sim \mathcal{N}(0,1)$
- Φ the CDF of Z

$$P(T > c) = P\left(\frac{T-\mu}{\sigma/\sqrt{n}} > \frac{c-\mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(Z > \frac{\sqrt{n}(c-\mu)}{\sigma}\right)$$

One Tailed Test Example

Israeltech challenge

- Z is standard normal - $Z \sim \mathcal{N}(0,1)$
- Φ the CDF of Z

$$
\begin{aligned}
P(T > c) &= P\left(\frac{T-\mu}{\sigma/\sqrt{n}} > \frac{c-\mu}{\sigma/\sqrt{n}}\right) \\
&= P\left(Z > \frac{\sqrt{n}(c-\mu)}{\sigma}\right) \\
&= 1 - \Phi\left(\frac{\sqrt{n}(c-\mu)}{\sigma}\right)
\end{aligned}
$$

One Tailed Test Example

$$P(T > c) = 1 - \Phi\left(\frac{\sqrt{n}\,(c-\mu)}{\sigma}\right)$$

- test is most strict for $\mu = 0$ why
- we require $P(T > c) = \alpha$

One Tailed Test Example

$$P\left(T > c\right) = 1 - \Phi\left(\frac{\sqrt{n}c}{\sigma}\right)$$

$$\alpha = 1 - \Phi\left(\frac{\sqrt{n}c}{\sigma}\right)$$

$$\frac{\sqrt{n}c}{\sigma} = \Phi^{-1}\left(1 - \alpha\right)$$

$$\boxed{c = \frac{\sigma\Phi^{-1}\left(1 - \alpha\right)}{\sqrt{n}}}$$

One Tailed Test Example

Israeltech challenge

$$P\left(T > c\right) = 1 - \Phi\left(\frac{\sqrt{n}c}{\sigma}\right)$$

$$\alpha = 1 - \Phi\left(\frac{\sqrt{n}c}{\sigma}\right)$$

$$\frac{\sqrt{n}c}{\sigma} = \Phi^{-1}\left(1 - \alpha\right)$$

$$\boxed{c = \frac{\sigma Z_\alpha}{\sqrt{n}}}$$

Two Tailed Test Example

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is known
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$

Two Tailed Test Example

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is known
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$
- **Test**: $R = \{X : T\left(X\right) > c\}$
- $T\left(X\right) = \frac{1}{n} \sum_{i=1}^{n} X_i$ ...

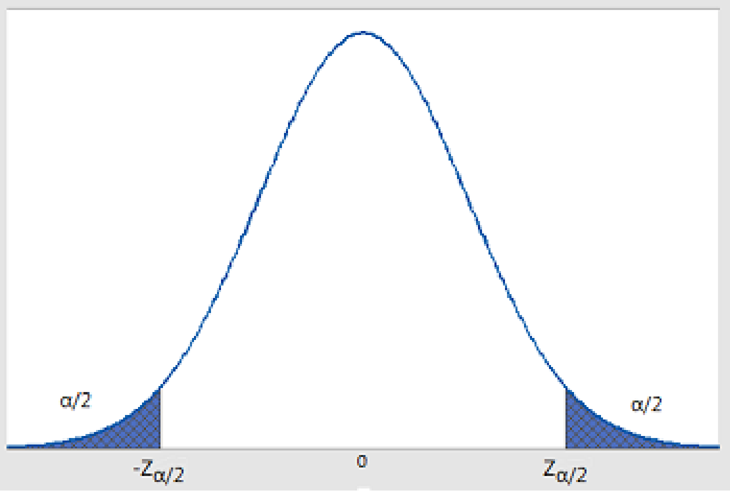
Two Tailed Test Example

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is known
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$
- **Test**: $R = \{X : T\left(X\right) > c\}$
- $T\left(X\right) = \frac{1}{n} \sum_{i=1}^{n} X_i$ ...
- nope

Two Tailed Test Example

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is known
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$
- **Test**: $R = \{X : |T(X)| > c\}$
- $T(X) = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}, \; T \sim \mathcal{N}(0, 1)$
- $c = Z_{\alpha/2}$

Two Tailed Test Example



z-test

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is $\boxed{\text{known}}$
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$
- **Test**: $R = \{X : |T(X)| > c\}$
- $T(X) = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$
- $T \sim \mathcal{N}(0, 1)$

z-test

Z-Test

a test where the test statistic $T \sim \mathcal{N}\left(0,1\right)$

what happens when σ is uknown?

t-test

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is $\boxed{\text{unknown}}$
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$
- $T\left(X\right) = \frac{\overline{X} - \mu_0}{\hat{\sigma}/\sqrt{n}}$

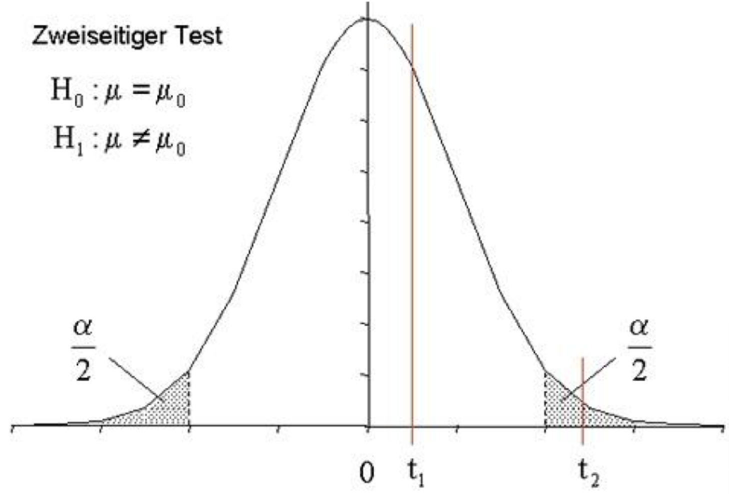
t-test

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is $\boxed{\text{unknown}}$
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$
- $T\left(X\right) = \frac{\overline{X} - \mu_0}{\hat{\sigma}/\sqrt{n}}$
- σ is estimated - sample standard deviation
- for large n, the distribution of T under H_0 tends to $\mathcal{N}\left(0, 1\right)$

t-test

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is $\boxed{\text{unknown}}$
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$
- $T\left(X\right) = \frac{\overline{X} - \mu_0}{\hat{\sigma}/\sqrt{n}}$
- σ is estimated - sample standard deviation
- exact distribution of T under H_0 is Student's t-distribution with $n - 1$ degrees of freedom

Student's t-distribution



Likelihood Ratio Test

- $X = X_1, \dots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is $\boxed{\text{known}}$
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$

Likelihood Ratio Test

Israeltech challenge

- $X = X_1, \ldots, X_n \sim \mathcal{N}\left(\mu, \sigma^2\right)$
- σ is $\boxed{\text{known}}$
- $H_0 : \mu = \mu_0$
- $H_1 : \mu \neq \mu_0$
- we can estimate μ_{mle}
- compute the likelihood ratio

rationale

$$\frac{\mathcal{L}\left(\mu_{mle}\right)}{\mathcal{L}\left(\mu_0\right)}$$

Likelihood Ratio Test

- test statistic

$$\lambda = 2\log\frac{\mathcal{L}\left(\theta_{mle}\right)}{\mathcal{L}\left(\theta_0\right)}$$

Likelihood Ratio Test

in our case

- $\mu_{mle} = \sum_{i=1}^{n} X_i/n$

$$\mathcal{L}(\mu) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2}$$

$$\lambda = 2\log\frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu_{mle})^2}}{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu_0)^2}}$$

Likelihood Ratio Test

$$\lambda = 2\log\frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu_{mle})^2}}{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu_0)^2}}$$

$$= 2\log e^{\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(X_i-\mu_0)^2-\sum_{i=1}^{n}(X_i-\mu_{mle})^2\right]}$$

$$\left[\begin{matrix}(a-b)^2\\ =\\ a^2-2ab+b^2\end{matrix}\right] = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(\mu_0^2-\mu_{mle}^2-2X_i\mu_0+2X_i\mu_{mle}\right)$$

$$= \frac{1}{\sigma^2}\left(n\left(\mu_0^2-\mu_{mle}^2\right)-2\left(\mu_0-\mu_{mle}\right)\sum_{i=1}^{n}X_i\right)$$

Likelihood Ratio Test

Israeltech challenge

$$\lambda = \frac{1}{\sigma^2}\left(n\left(\mu_0^2 - \mu_{mle}^2\right) - 2\left(\mu_0 - \mu_{mle}\right)\frac{n}{n}\sum_{i=1}^{n} X_i\right)$$

$$\left[\begin{matrix}\left(a^2-b^2\right) \\ = \\ (a+b)(a-b)\end{matrix}\right] = \frac{1}{\sigma^2}\left(\begin{matrix} n\left(\mu_0 - \mu_{mle}\right)\left(\mu_0 + \mu_{mle}\right) \\ - \\ 2n\left(\mu_0 - \mu_{mle}\right)\mu_{mle}\end{matrix}\right)$$

$$= \frac{n\left(\mu_0 - \mu_{mle}\right)^2}{\sigma^2}$$

Likelihood Ratio Test

$$\lambda = \frac{n\left(\mu_0 - \mu_{mle}\right)^2}{\sigma^2}$$

- under H_0

Likelihood Ratio Test

Israeltech challenge

$$\lambda = \frac{n\left(\mu_0 - \mu_{mle}\right)^2}{\sigma^2}$$

- under H_0

$$\mu_{mle} \sim \mathcal{N}\left(\mu_0, \sigma^2/n\right)$$

Likelihood Ratio Test

$$\lambda = \frac{n\left(\mu_0 - \mu_{mle}\right)^2}{\sigma^2}$$

- under H_0

$$\mu_{mle} \sim \mathcal{N}\left(\mu_0, \sigma^2/n\right)$$

$$\sqrt{\lambda} = \frac{\sqrt{n}}{\sigma}\left(\mu_0 - \mu_{mle}\right) \sim \mathcal{N}\left(0, 1\right)$$

Likelihood Ratio Test

$$\lambda = \frac{n\left(\mu_0 - \mu_{mle}\right)^2}{\sigma^2}$$

- under H_0

$$\mu_{mle} \sim \mathcal{N}\left(\mu_0, \sigma^2/n\right)$$

$$\sqrt{\lambda} = \frac{\sqrt{n}}{\sigma}\left(\mu_0 - \mu_{mle}\right) \sim \mathcal{N}\left(0, 1\right)$$

$$\boxed{\lambda \sim \chi_1^2}$$

Likelihood Ratio Test

IsraelTech challenge

$$\lambda = \frac{n\left(\mu_0 - \mu_{mle}\right)^2}{\sigma^2}$$

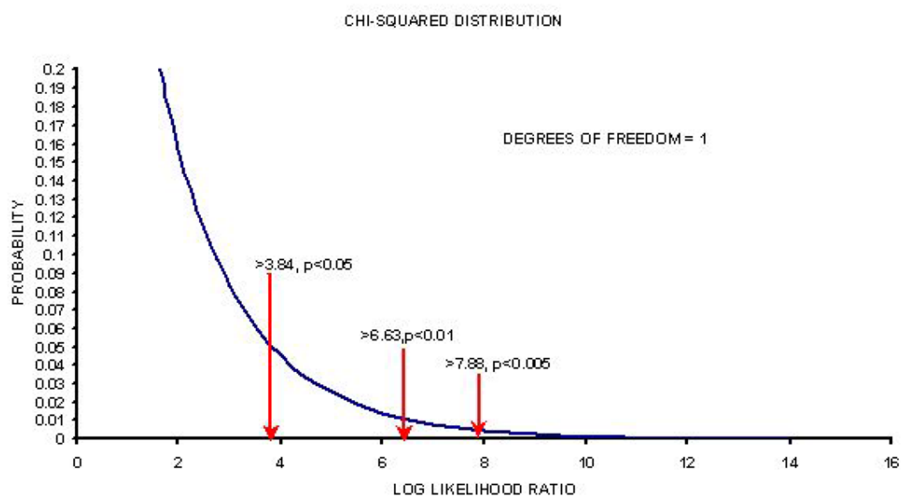
- under H_0

$$\mu_{mle} \sim \mathcal{N}\left(\mu_0, \sigma^2/n\right)$$

$$\boxed{\lambda \sim \chi_1^2}$$

- The Chi-Distribution with 1 degree of freedom
- asymptotically true for a wide range of cases

Chi Squared Distribution



Testing Independence

X and Y two random variables

- let's test independence

Testing Independence

X and Y two random variables

- H_0 X and Y are independent
- H_1 X and Y are **not** independent

Testing Independence

under H_0 what can we say about $P\left(X,Y\right)$?

Testing Independence

under H_0 what can we say about $P\left(X,Y\right)$?

$$P\left(X,Y\right)=P\left(X\right)P\left(Y\right)$$

Testing Independence

IsraelTech challenge

- X and Y are binary
- $(X_1, Y_1), \ldots, (X_n, Y_n)$ samples of (X, Y)
- number of observations for each value

	$Y = 0$	$Y = 1$	Total
$X = 0$	z_{00}	z_{01}	$z_{0\cdot}$
$X = 1$	z_{10}	z_{11}	$z_{1\cdot}$
Total	$z_{\cdot 0}$	$z_{\cdot 1}$	n

Testing Independence

Israeltech challenge

- X and Y are binary
- $(X_1, Y_1), \ldots, (X_n, Y_n)$ samples of (X, Y)
- probabilities

	$Y = 0$	$Y = 1$	Total
$X = 0$	p_{00}	p_{01}	$p_{0\cdot}$
$X = 1$	p_{10}	p_{11}	$p_{1\cdot}$
Total	$p_{\cdot 0}$	$p_{\cdot 1}$	1

Testing Independence

estimation

$$\hat{p}_{ij} = \frac{z_{ij}}{n}$$
$$\hat{p}_{i\cdot} = \frac{z_{i\cdot}}{n}$$
$$\hat{p}_{\cdot j} = \frac{z_{\cdot j}}{n}$$

Testing Independence

example

we sample (X, Y) $n = 3$ times

$$(0,1), (0,0), (1,0)$$

number of observations for each value

	$Y=0$	$Y=1$	Total
$X=0$	1	1	2
$X=1$	1	0	1
Total	2	1	3

Testing Independence

example

we sample (X, Y) $n = 3$ times

$$(0,1), (0,0), (1,0)$$

our estimation of the probabilities

	$Y = 0$	$Y = 1$	Total
$X = 0$	$1/3$	$1/3$	$2/3$
$X = 1$	$1/3$	0	$1/3$
Total	$2/3$	$1/3$	1

Testing Independence

example

we sample (X, Y) $n = 3$ times

$$(0,1),(0,0),(1,0)$$

$$P(X=0, Y=1) = ?$$

- under H_0

$$\begin{aligned} P(X=0, Y=1) &= P(X=0)\,P(Y=1) \\ &= 2/3 \cdot 1/3 \\ &= 2/9 \end{aligned}$$

- under H_1

$$P(X=0, Y=1) = 1/3$$

Testing Independence

IsraelTech challenge

$$\begin{aligned}
\lambda &= 2\log\frac{\mathcal{L}\left(\hat{p}\right)}{\mathcal{L}\left(\hat{p}_0\right)} \\
&= 2\sum_{i=1}^{n}\log\frac{\hat{p}_{X_iY_i}}{\hat{p}_{X_i\cdot}\hat{p}_{\cdot Y_i}} \\
&= 2\sum_{i=0}^{1}\sum_{j=0}^{1} z_{ij}\log\frac{\hat{p}_{ij}}{\hat{p}_{i\cdot}\hat{p}_{\cdot j}} \\
&= 2\sum_{i=0}^{1}\sum_{j=0}^{1} z_{ij}\log\frac{z_{ij}n}{z_{i\cdot}z_{\cdot j}}
\end{aligned}$$

Testing Independence

$$\lambda = 2 \sum_{i=0}^{1} \sum_{j=0}^{1} z_{ij} \log \frac{z_{ij} n}{z_{i.} z_{.j}}$$

- under H_0, $\lambda \sim \chi^2_1$

Testing Independence

In the general case

- X - r values observed
- Y - c values observed

$$\lambda = 2 \sum_{i=0}^{1} \sum_{j=0}^{1} z_{ij} \log \frac{z_{ij} n}{z_{i.} z_{.j}}$$

- under H_0, $\lambda \sim \chi_k^2$
- $k = (r-1)(c-1)$

Testing Independence

Pearson's Test

- $U = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(z_{ij} - z_{i\cdot} z_{\cdot j}/n)^2}{z_{i\cdot} z_{\cdot j}/n}$
- under H_0, $U \sim \chi^2_{(r-1)(c-1)}$

Testing Correlation

- X, Y two random variables
- $(X_1, Y_1,), \ldots, (X_n, Y_n)$ samples of (X, Y)

how should we test?

Testing Correlation

- X, Y two random variables
- $(X_1, Y_1,), \ldots, (X_n, Y_n)$ samples of (X, Y)
- H_0 - $\rho_{XY} = 0$
- H_1 - $\rho_{XY} \neq 0$
- r the sample correlation

Testing Correlation

- no knowledge of $P(r)$ under H_0
- maybe we need a different approach

Testing Correlation

- let's generate an empiric $P(r)$
- p a random permutation of $1, \ldots, n$

$$p : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$$

- example

$$p(\{1, 2, 3\}) = \{3, 1, 2\}$$

- a "new" (permuted) set of samples

$$\left(X_1, Y_{p(1)}\right), \ldots, \left(X_1, Y_{p(n)}\right)$$

Testing Correlation

- generate $p_1, \ldots, p_m$
- $\forall j = 1, \ldots, m$ compute r_j from

$$\left(X_1, Y_{p_j(1)}\right), \ldots, \left(X_1, Y_{p_j(n)}\right)$$

- the P-Value for this empiric distribution

$$P_{val} = \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}\left\{|r_j| > |r|\right\}$$

Important Terms

- **power/sensitivity/recall** - probability of correctly rejecting H_0 when H_1 is true

$$P\left(\text{reject } H_0 \mid H_1\right)$$

- **specificity** - probability of correctly not rejecting H_0 when it is true

$$P\left(\text{not reject } H_0 \mid H_0\right)$$

- **precision** - probability of a detection being true

$$P\left(H_1 \mid \text{reject } H_0\right)$$

Conclusion

- from P-Values to test statistics
- one and two tailed tests
- the statistical test zoo
- tests for specific cases of interest - independence, correlation

Credits

figures

- two tailed test - onlinecourses.science.psu.edu/stat500/node/44