Advanced statistical methods

Frédéric Pascal

CentraleSupélec, Laboratory of Signals and Systems (L2S), France
frederic.pascal@centralesupelec.fr
http://fredericpascal.blogspot.fr

MSc in Data Sciences & Business Analytics
CentraleSupélec / ESSEC
Oct. 2^{nd} - Dec. 20^{th}, 2017

Part C

Hypothesis testing - Detection theory

Part C: Contents

Key references of Part C

From an EE / SP point of view...

- Kay, Steven M. *Fundamentals of Statistical Signal Processing - Detection Theory*, Vol. 2, Prentice Hall, 1998.
- Poor, Vincent, H. *An Introduction to Signal Detection and Estimation*, 2nd ed, Springer, 1998.

From a statistical point of view...

- Lehmann, Erich L., and Romano, Joseph P. *Testing Statistical Hypotheses*, Springer, 2006.
- Casella, George, and Roger L. Berger. *Statistical inference*, Vol. 2. Pacific Grove, CA: Duxbury, 2002.

+ many many references...

I. Generalities

- Principles
- Errors, power and level of a test
- Neyman approach

II. UMP tests

III. Student-t test

IV. Asymptotic Tests

Generalities

Let a n-sample $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ i.i.d. $\sim P_\theta, \theta \in \Theta$. Let H_0 and H_1, 2 non-empty disjoint subsets of Θ s.t. $H_0 \cup H_1 = \Theta$.

H_0 is the null hypothesis while H_1 is called the alternative hypothesis.

Remember: no symmetry!

Goal: To find a procedure that allows to decide whether θ belongs to H_0 or not, regarding the datasets $x = (x_1, \ldots, x_n) \in \mathscr{X}^n$.

Definition

An hypothesis is said ***simple*** *if it is reduced to a single element. Else, it is called* ***composite****.*

Definition

A ***(pure) test*** *is a mapping δ from $\mathscr{X}^n$ onto $\{0, 1\}$ s.t.:*
If $\delta(x) = 0$, one decides H_0, while if $\delta(x) = 1$, one rejects H_0.
The region $W = \{x \in \mathscr{X}^n \mid \delta(x) = 1\}$ is called the ***rejection region*** *or the* ***critical region****. Its complement is called the* ***acceptance region****.*

Generalities

Remark

A test is characterized (and will be identified) by its rejection region W.

Definition (Different errors)

For a test, there are two possible errors:

- *rejecting H_0 when it is true: **type-I error or error of 1^{st} kind**.*
- *accepting H_0 when it is false: **type-II error or error of 2^{nd} kind**.*

Definition (Type-I and Type-II errors)

For a test δ with critical region W, one has

- ***Type-I error:*** $\alpha_W : \begin{cases} H_0 \to [0,1] \\ \theta \mapsto P_\theta(W); \end{cases}$
- ***Type-II error:*** $\beta_W : \begin{cases} H_1 \to [0,1] \\ \theta \mapsto P_\theta(W^c) = 1 - P_\theta(W). \end{cases}$

Generalities

Definition (Power of the test)

The ***power*** *of a test W is defined as:*

$$\rho_W : \begin{cases} H_1 \to [0,1] \\ \theta \mapsto P_\theta(W) = 1 - \beta_W(\theta). \end{cases}$$

Definition (Randomized test (more general))

A random test is a mapping φ from $\mathscr{X}^n$ into $[0,1]$ where $\varphi(x)$ is the probability of rejecting H_0 for the dataset $x = (x_1, \cdots, x_n) \in \mathscr{X}^n$.

Remark

For $\varphi = \mathbb{1}_W$, one retrieves the simple test!

Generalities

Definition (Type-I and Type-II errors, power for a test φ)

- ***Type-I error:*** $\alpha_\varphi : \begin{cases} H_0 \to [0,1] \\ \theta \mapsto E_\theta\left[\varphi(\mathbf{x})\right]; \end{cases}$
- ***Type-II error:*** $\beta_\varphi : \begin{cases} H_1 \to [0,1] \\ \theta \mapsto 1 - E_\theta\left[\varphi(\mathbf{x})\right]; \end{cases}$
- ***Power of the test:*** $\rho_\varphi = 1 - \beta_\varphi = E_{H_1}\left[\varphi(\mathbf{x})\right].$

Definition (Level of significance (ls))

*The **level of significance** α (typically 0.01 or 0.05 as for the IC) for a test φ is:*

$$\alpha = \sup_{\theta \in H_0} \alpha_\varphi(\theta) = \sup_{\theta \in H_0} E_\theta\left[\varphi(\mathbf{x})\right].$$

Neyman Principle

Goal: one wants to control (or fix) the type-I error, i.e. the probability of rejecting H_0 when it is true.

The Neyman principle consists in considering all tests with a ls ≤ to a fixed α, and then, in finding (among these tests) the one with the smallest Type-II error.

Since $\rho_\varphi = 1 - \beta_\varphi$, such test will said to be UMP.

Definition (Uniformly Most Powerful (UMP))

φ is UMP at the threshold α if its ls $\leq \alpha$ and if $\forall \varphi'$ with a ls $\leq \alpha$, one has:

$$\forall \theta \in H_1, E_\theta\left[\varphi(\mathbf{x})\right] \geq E_\theta\left[\varphi'(\mathbf{x})\right].$$

I. Generalities

II. UMP tests

- Simple hypothesis testing
- Composite tests - One-sided hypotheses

III. Student-t test

IV. Asymptotic Tests

Simple hypothesis testing

In this part, for the n-sample $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, one considers,

$$H_0 : \{\theta = \theta_0\} \text{ versus } H_1 : \{\theta = \theta_1\},$$

which means that $\Theta = \{\theta_0, \theta_1\}$.

So, 2 probabilities P_{θ_0} (or P_0) and P_{θ_1} (or P_1), that implies 2 LF $L_0(x) = L(x;\theta_0)$ and $L_1(x) = L(x;\theta_1)$, for $x = (x_1, \ldots, x_n) \in \mathscr{X}^n$.

Definition (Neyman test or Likelihood Ratio Test (LRT))

A Neyman test is a test φ s.t. $\exists k \in \mathbb{R}_+^*$, and

$$\varphi(x) = \begin{cases} 1 & \text{if} \quad L(x;\theta_1) > k\, L(x;\theta_0) \\ 0 & \text{if} \quad L(x;\theta_1) < k\, L(x;\theta_0) \end{cases}$$

The value of φ is not specified for $\left\{x \in \mathscr{X}^n | L_1(x) = k\, L_0(x)\right\}$.

Neyman-Pearson Lemma

Remark

$L_1(x)/L_0(x)$ *is called the* ***Likelihood Ratio (LR)****. The Neyman test consists in accepting the most likely hypothesis for a given observation* x.

Proposition (Neyman-Pearson Lemma)

1. ***Existence*** $\forall \alpha \in (0,1)$, *it exists a Neyman test s.t.* $E_{\theta_0}(\varphi) = \alpha$. *Moreover,* k *is the quantile of order* $(1-\alpha)$ *of the LR distribution* $\frac{L_1(x)}{L_0(x)}$ *under* P_0 *and one can impose that* φ *is constant for* $x \in \mathscr{X}^n$ *s.t.* $L_1(x) = kL_0(x)$. *If the LR CDF under* P_0 *evaluated in* k *is* $(1-\alpha)$ *(continuous CDF), thus one can choose this constant* $= 0$ *(pure test).*
2. ***S. cond.*** $\forall \alpha \in (0,1)$, *a Neyman test s.t.* $E_{\theta_0}(\varphi) = \alpha$ *is* ***UMP*** *at level* α.
3. ***N. cond.*** $\forall \alpha \in (0,1)$, *a UMP test at level* α *is* ***necessarily a Neyman test****.*

Proof

Essential to built the Neyman test...

Neyman-Pearson Lemma

Remark

1. *Conclusion: the only UMP tests at level α are the Neyman tests of level of significance α.*
2. *If the LR CDF under H_0 is continuous, one obtains the test of critical region $W = \left\{x \in \mathscr{X}^n \mid L_1(x) > k\,L_0(x)\right\}$ where k is defined by $P_0(L_1(X) > k\,L_0(X)) = \alpha$.*
3. *The power $E_1(\varphi)$ of a UMP test at level α is necessarily $\geq \alpha$. Indeed, φ is preferable to the constant test $\psi = \alpha$ (which is of ls α), thus $E_1(\varphi) \geq E_1(\psi) = \alpha$.*

Neyman-Pearson Lemma

Example 1: Let us consider the exponential model (1)

$$L(x,\theta) = C(\theta)\,h(x)\exp\left[\sum_{j=1}^{d} Q_j(\theta)\,S_j(x)\right]$$

where $\theta \in \{\theta_0, \theta_1\}$, with $\theta_1 > \theta_0$. Assume an identifiable model: $Q(\theta_0) \neq Q(\theta_1)$ $\big($e.g., $Q(\theta_1) > Q(\theta_0)\big)$.
Goal: test $H_0 : \{\theta = \theta_0\}$ versus $H_1 : \{\theta = \theta_1\}$.

Example 2: Let us consider $(X_1, \cdots, X_n) \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ with σ^2 known.
Goal: test $H_0 : \{\mu = \mu_0\}$ versus $H_1 : \{\mu = \mu_1\}$, with $\mu_0 < \mu_1$.

Example 3: Let us consider $(X_1, \cdots, X_n) \overset{iid}{\sim} Poisson(\theta)$.
Goal: test $H_0 : \{\theta = \theta_0\}$ versus $H_1 : \{\theta = \theta_1\}$, with $\theta_0 < \theta_1$.

Composite tests - One-sided hypotheses

Now, let us consider a model with only 1 parameter and where Θ is an interval of $\mathbb{R}$. One assume $L(x,\theta) > 0, \forall x \in \mathscr{X}^n, \forall \theta \in \Theta$.

Goal: **test** $H_0 : \{\theta \leq \theta_0\}$ **versus** $H_1 : \{\theta > \theta_0\}$.

More general problem!

Let us consider the family having **monotone likelihood ratio**:

Definition (Monotone LR)

The family $\{P_\theta^{\otimes n}, \theta \in \Theta\}$ is said to have ***monotone likelihood ratio*** *if it exists a real-valued statistic $U(x)$ s.t. $\forall \theta' < \theta''$, $\frac{L(x,\theta'')}{L(x,\theta')}$ is a strictly increasing (or decreasing) function of U.*

Remark

By changing U into $-U$, one can always assume strictly increasing in previous definition.

Lehman Theorem

Theorem (Lehman theorem)

Let $\alpha \in (0,1)$. If the family $(P_\theta, \theta \in \Theta)$ has ***monotone (increasing) likelihood ratio****, there exists a UMP test at level α for testing $H_0 : \{\theta \leq \theta_0\}$ versus $H_1 = \{\theta > \theta_0\}$. This test is defined by:*

$$\begin{cases} \varphi(x) = 1 & \text{if} \quad U(x) > c \\ \varphi(x) = \gamma & \text{if} \quad U(x) = c \\ \varphi(x) = 0 & \text{if} \quad U(x) < c \end{cases}$$

where c and γ are obtained with $E_{\theta_0}[\varphi] = \alpha$. The same test is UMP at level α for testing:

1. $H_0 : \{\theta = \theta_0\}$ *versus* $H_1 : \{\theta > \theta_0\}$
2. $H_0 : \{\theta = \theta_0\}$ *versus* $H_1 : \{\theta = \theta_1\}$

where $\theta_1 > \theta_0$.

Lehman Theorem

Remark

If the inequalities are reversed in the test, i.e. $H_0 : \{\theta \geq \theta_0\}$ and $H_1 : \{\theta < \theta_0\}$, then the UMP test is obtained by reversing the inequalities (in the test).

Example: The exponential model with LF $L(x,\theta) = C(\theta)\, h(x) \exp\big(Q(\theta)\, S(x)\big)$ where $Q(\theta)$ is strictly increasing, has increasing LR with $U(X) = S(X)$.

Remark (Important)

In general, it does NOT exist UMP test for testing $H_0 : \{\theta = \theta_0\}$ versus $H_1 : \{\theta \neq \theta_0\}$ (even for monotone LR).
For instance, let's consider the Gaussian model, σ^2 known. The UMP test for $H_0 : \{\mu = \mu_0\}$ versus $H_1 : \{\mu > \mu_0\}$ is $\begin{cases} \rho(x) = 1 & \text{if } \sum x_i > c \\ \rho(x) = 0 & \text{if } \sum x_i \leq c \end{cases}$ *while the UMP test for $H_0 : \{\mu = \mu_0\}$ versus $H_1 : \{\mu < \mu_0\}$ is* $\begin{cases} \rho(x) = 1 & \text{if } \sum x_i < c \\ \rho(x) = 0 & \text{if } \sum x_i \geq c \end{cases}$

$\Rightarrow$ ***no UMP test for testing $\mu = \mu_0$ versus $\mu \neq \mu_0$.***

I. Generalities

II. UMP tests

III. Student-t test

IV. Asymptotic Tests

Student test

Let $(X_1, \cdots, X_n) \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ with μ and σ^2 unknown.

Goal: test $H_0 : \{\mu = \mu_0\}$ versus $H_1 : \{\mu \neq \mu_0\}$ at level $\alpha \in (0,1)$.

General methodology

1. From the *Student theorem*, one has

$$T_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \sim t(n-1)$$

   where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$.
2. Under H_0:

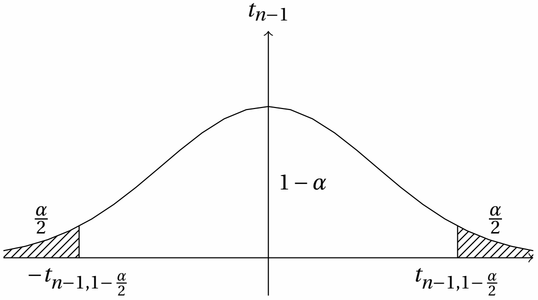
$$\xi_n = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{S_n} \sim t(n-1)$$

3. Under H_1: From the SLLN, $\bar{X}_n - \mu_0 \xrightarrow[n\to\infty]{a.s} \mu - \mu_0$ and $S_n \xrightarrow[n\to\infty]{a.s} \sigma$.
   Thus $\xi \xrightarrow[n\to\infty]{a.s} +\infty$ if $\mu > \mu_0$ and $\xi \xrightarrow[n\to\infty]{a.s} -\infty$ if $\mu < \mu_0$
4. Critical region:

$$\mathbf{W_n} = \{|\xi_\mathbf{n}| > \mathbf{a}\}$$

Student test

- Let $t_{n-1,r}$ the quantile of order r of the t-distribution t_{n-1}:



Thus, under $H_0, P(|\xi_n| > t_{n-1,1-\frac{\alpha}{2}}) = \alpha$.

Previously, one have seen that $I_n = \left[\bar{X}_n - \frac{t_{n-1,1-\alpha/2}S_n}{\sqrt{n}}, \bar{X}_n + \frac{t_{n-1,1-\alpha/2}S_n}{\sqrt{n}}\right]$ is a $(1-\alpha)$-CI for μ_0. Here is the link between CI and Student (bilateral) test $\mu_0 \in I_n$ iff $|\xi_n| \leq t_{n-1,1-\frac{\alpha}{2}}$. Finally, the associated p-value is $p = P(|T| > |\xi_n^{obs}|)$ where $T \sim t(n-1)$ and ξ_n^{obs} is the observed value of ξ_n.

I. Generalities

II. UMP tests

III. Student-t test

IV. Asymptotic Tests

- Generalities
- Wald test
- Rao (score) test and LRT
- χ^2 tests

Generalities

As for estimators, in many situations, one CANNOT find the distribution of the LR (or the statistic of the monotone LR). As a consequence, one cannot set the parameters k and γ for the test.

A solution (like in point estimation theory) is to rely on asymptotic properties!

Now, instead of considering a test W, we will consider a sequence of tests $(W_n)_{n\in\mathbb{N}^*}$.

Definition (Asymptotic level)

*An asymptotic test W_n is at **asymptotic level** α if*

$$\lim_{n\to\infty} \sup_{\theta\in H_0} P_\theta(W_n) = \alpha\,.$$

Generalities

Definition (Uniform asymptotic level)

*An asymptotic test W_n is at **uniform asymptotic level** α if*

$$\sup_{\theta \in H_0} \lim_{n \to \infty} P_\theta(W_n) = \alpha \,.$$

Definition (Consistant (or convergent) test)

*An asymptotic test W_n is said to be **consistant (or convergent)** if its power tends towards 1, i.e.,*

$$\forall \theta \in H_1 \,, \lim_{n \to \infty} P_\theta(W_n) = 1 \,.$$

This means that the Type-II error tends to 0!

Example: the t-test is consistant...

Asymptotic tests

Implicit constraint: $H_0 : \left\{\theta | g(\theta) = 0\right\}$

where g a mapping from $\mathbb{R}^d$ into $\mathbb{R}^r$, of class C^1 s.t. the $r \times d$ matrix

$$\frac{\partial g}{\partial \theta^t} = \left(\frac{\partial g_i}{\partial \theta_j}\right)_{1 \leq i \leq r, 1 \leq j \leq d} \text{ is of rank } r \text{ (so } r \leq d).$$

Goal: test $H_0 : \{\theta \in \Theta, g(\theta) = 0\}$ versus the alternative hypothesis $H_1 : \{\theta \in \Theta, g(\theta) \neq 0\}$

More general than $H_0 : \{\theta = \theta_0\}$ versus $H_1 : \{\theta \neq \theta_0\}$

To answer such problems, there exist (at least) 3 asymptotic tests:

- Wald test
- Rao (score) test
- Likelihood Ratio Test (LRT)

Wald test

Proposition (Wald test)

Let $\hat{\theta}_n^{ML}$ the MLE of θ. Under H_0, the sequence of r.V., one has: $\left(\sqrt{n} g(\hat{\theta}_n^{ML})\right) \xrightarrow[n\to\infty]{dist.} \mathcal{N}\left(\mathbf{0}, \Sigma(\theta_0)\right)$, where $\theta_0 \in H_0$ is the true value of the parameter θ and where $\Sigma(\theta_0) = \dfrac{\partial g}{\partial \theta^t}(\theta_0) I_1(\theta_0)^{-1} \dfrac{\partial g^t}{\partial \theta}(\theta_0)$.

Furthermore, the test statistic $\xi_n^W = n\, g\left(\hat{\theta}_n^{ML}\right)^t \Sigma\left(\hat{\theta}_n^{ML}\right)^{-1} g\left(\hat{\theta}_n^{ML}\right)$ converges in distribution under H_0 towards a χ^2-distribution with r d.o.f.:

$$\xi_n^W \xrightarrow[n\to\infty]{dist.} \chi^2(r)$$

The Wald tests are defined by the following critical region:

$$W_n = \left\{\xi_n^W > q_r(1-\alpha)\right\}$$

where $q_r(1-\alpha)$ is the quantile of order $(1-\alpha)$ of the χ^2-distribution with r d.o.f. This test is strongly convergent at asymptotic level $\alpha = P(\chi^2(r) > q_r(1-\alpha))$.

Wald test

Definition (p-value)

The asymptotic p-value of the Wald test is defined by

$$p = P\left(\chi^2(r) > \xi_n^W(x_1,\ldots,x_n)\right)$$

where $\chi^2(r)$ is a r.v. following a χ^2-dist. with r d.o.f. and $\xi_n^W(x_1,\ldots,x_n)$ is the observed test statistic. One rejects H_0 if $p<\alpha$...

Remark

If one cannot compute $I_1(\theta)$. One can estimate $I_1(\theta)$ by the MM and replace it in the Wald test WITHOUT changing the results!:

$$\hat{I}_1(\cdot) = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \ln L(x_i,\cdot)}{\partial \theta^t}\frac{\partial \ln L(x_i,\cdot)^t}{\partial \theta} \quad \text{ou} \quad \hat{I}_1(\cdot) = -\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln L(x_i,\cdot)}{\partial \theta \partial \theta^t}.$$

Proof (Wald test)

Allows to understand the methodology...

Wald test

Example: Let a Gaussian n-sample $\begin{pmatrix} X_i \\ Y_i \end{pmatrix}_{i\in\{1,\ldots,n\}} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}\right)$ with σ_1 and σ_2 known. Let $\theta = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$.

Goal: test $\mu_1 = \mu_2$, i.e., $H_0 : \{\mu_1 - \mu_2 = 0\}$ versus $H_1 : \{\mu_1 - \mu_2 \neq 0\}$.

Let us set $g(\theta) = \mu_2 - \mu_1$ and show that the Wald test statistic is

$$\xi_n^W = \frac{n(\hat{\mu}_1 - \hat{\mu}_2)^2}{\sigma_1^2 + \sigma_2^2}$$

where $\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n} Y_i$. One has

$$\xi_n^W \xrightarrow[n\to\infty]{dist.} \chi^2(1)$$

Rao-score test and Likelihood Ratio test (LRT)

Let $\hat{\theta}_n^c$ the MLE of θ under the constraint $g(\theta) = 0$, i.e. under H_0.

Theorem (**Rao test and LRT**)

The test statistics are defined by:

$$\xi_n^R = \frac{1}{n}\frac{\partial \ln L(x_i,\ldots,x_n;\hat{\theta}_n^c)}{\partial \theta^t} I_1(\hat{\theta}_n^c)^{-1}\frac{\partial \ln L(x_i,\ldots,x_n;\hat{\theta}_n^c)^t}{\partial \theta}$$

$$\xi_n^{LR} = 2(\ln L(x_i,\ldots,x_n;\hat{\theta}_n) - \ln L(x_i,\ldots,x_n;\hat{\theta}_n^c))$$

Rao test and the LRT are defined by the following critical region

$$W_n = \{\xi_n^i > q_r(1-\alpha)\}$$

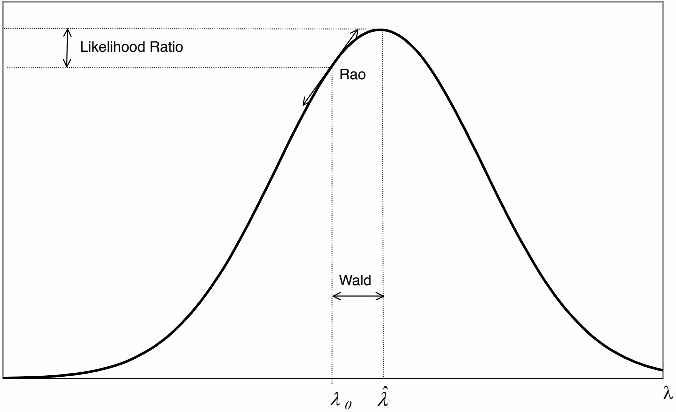
where $q_r(1-\alpha)$ is the quantile of order $(1-\alpha)$ of the χ^2-distribution with r d.o.f. These tests are strongly convergent at asymptotic level $\alpha = P(\chi^2(r) > q_r(1-\alpha))$. Furthermore, under H_0, one has:

$$\xi_n^W - \xi_n^R \xrightarrow[n\to\infty]{P} 0 \text{ and } \xi_n^W - \xi_n^{LR} \xrightarrow[n\to\infty]{P} 0$$

Rao-score test and Likelihood Ratio test (LRT)

Example Testing $H_0 : \{\lambda = \lambda_0\}$ versus $H_1 : \{\lambda \neq \lambda_0\}$ in case of a Poisson distribution with parameter λ...

...



χ^2 test: Goodness-of-Fit to a given distribution

<u>Goal:</u> test the goodness of fit of r.V. to a discrete and finite distribution (e.g., binomial, ...)

Quite restrictive but it CAN be extended to all distributions!

Let the n-sample $(X_1, \ldots, X_n)$ i.i.d. with values in $\{a_1, \cdots, a_m\}$ and distribution P, where P is characterized by its weights $P = (p_1, \cdots p_m)$ (it is a PMF) with $\sum_{i=1}^{m} p_i = 1$ and $\forall j = 1, \ldots, n, \forall i = 1, \ldots, m, p_i = P(X_j = a_i)$.

One wants to test $H_0 : \{P = P_{p_0}\}$, where $p_0 = (p_1^0, \cdots, p_m^0)$ **is given** (no unknown parameter) with $\sum_{i=1}^{m} p_i^0 = 1, p_i^0 > 0, \forall i = 1, \ldots, m$.

Let N_i the counting statistic and p_i is the empirical frequency of $\{X_k = a_i\}$:

$$N_i = \sum_{k=1}^{n} \mathbb{1}_{\{X_k = a_i\}} \quad \text{and} \quad \hat{p}_i = \frac{N_i}{n}$$

χ^2 test: Goodness-of-Fit to a given distribution

Theorem (χ^2- test)

Under H_0

$$\xi_n = \sum_{i=1}^{m} \frac{(N_i - np_i^0)^2}{np_i^0} = n \sum_{i=1}^{m} \frac{(\hat{p}_i - p_i^0)^2}{p_i^0}$$

And ξ_n converges in distribution towards a χ^2-distribution with $(m-1)$ d.o.f. when $n \to +\infty$.
The test is defined by the critical region:

$$W_n = \{\xi_n > q_{m-1}(1-\alpha)\}$$

where $q_{m-1}(1-\alpha)$ is the quantile of order $(1-\alpha)$ of the χ^2-distribution with $(m-1)$ d.o.f. This test is strongly convergent at asymptotic level $\alpha = P(\chi^2(m-1) > q_{m-1}(1-\alpha))$.

Example: Toss a coin...

χ^2 test: Goodness-of-Fit to a given distribution

Now, let us test $H_0 : \{p = p(\theta)\}$ versus $H_1 : \{p \neq p(\theta)\}$ where $\theta \in \Theta \subset \mathbb{R}^d$, Θ open-set and θ is unknown!

Theorem (General χ^2- test)

Under H_0

$$\xi_n = \sum_{i=1}^{m} \frac{(N_i - np_i(\hat{\theta}_n))^2}{n\,p_i(\hat{\theta}_n)} = n \sum_{i=1}^{m} \frac{(\hat{p}_i - p_i(\hat{\theta}_n))^2}{p_i(\hat{\theta}_n)}$$

where $\hat{\theta}_n$ is the MLE of θ.
And ξ_n converges in distribution towards a χ^2-distribution with $(m-1-d)$ d.o.f. when $n \to +\infty$.
The test is defined by the critical region:

$$W_n = \{\xi_n > q_{m-1-d}(1-\alpha)\}$$

where $q_{m-1-d}(1-\alpha)$ is the quantile of order $(1-\alpha)$ of the χ^2-distribution with $(m-1-d)$ d.o.f. This test is strongly convergent at asymptotic level $\alpha = P(\chi^2(m-1-d) > q_{m-1-d}(1-\alpha))$.

χ^2 test: Goodness-of-Fit to a given distribution

How to generalized those χ^2 tests to continuous distribution or infinite discrete distribution?

Remark (On the use of χ^2 tests!)

- *It is an asymptotic test. In practice, it works if $n\, p_i(\hat{\theta}_n) > 5, \forall i$ and if $N_i \geq 5, \forall i$. Else, one regroups classes (cf exercise in the problems).*
- *In case of continuous r.v. with unknown distribution, one wants to test if it belongs to the family $\{P_\theta, \theta \in \Theta\}$. The idea is to partition $\mathbb{R}$ into m intervals $(A_i)_{i=1,\dots,m}$. The choice of m is a tradeoff:*
  - *m should be sufficiently large so that the discrete dist. $\{\pi_i = \pi(A_i)\}$ and $\{p_{\theta,i} = P_\theta(A_i)\}$ be sufficiently close to π and P_θ (if m is small, the test will be less powerful).*
  - *One the other hand, m should not be too large so that the $p_{\theta,i}$ be sufficiently large to satisfy $np_i(\hat{\theta}_n) > 5$.*

χ^2 test for independence

Let $(X_k, Y_k), k = 1, ..., n$ i.i.d. with values in $\{a_1, \cdots, a_l\} \times \{b_1, \cdots, b_r\}$. Let us denote $p_{i,j} = P(X_1 = a_i, Y_1 = b_j)$ and

$$p_{i,.} = P(X_1 = a_i) = \sum_{j=1}^{r} p_{i,j} \text{ and } p_{.,j} = P(Y_1 = b_j) = \sum_{i=1}^{l} p_{i,j}$$

One wants to know if X_1 and Y_1 are independent, i.e. if

$$H_0 : \{p_{i,j} = p_{i,.}.p_{.,j}, \forall\, i, j\}$$

Let $N_{i,j} = \sum_{k=1}^{n} \mathbb{1}_{\{X_k = a_i, Y_k = b_j\}}$ the counting statistic and

$$N_{i,.} = \sum_{k=1}^{n} \mathbb{1}_{\{X_k = a_i\}} \text{ and } N_{.,j} = \sum_{k=1}^{n} \mathbb{1}_{\{Y_k = b_j\}}$$

χ^2 test for independence

Theorem (χ^2- test for independence)

Under H_0

$$\xi_n = \sum_{i=1}^{l} \sum_{j=1}^{r} \frac{\left(N_{i,j} - \frac{N_{i,\cdot} N_{\cdot,j}}{n}\right)^2}{\frac{N_{i,\cdot} N_{\cdot,j}}{n}}$$

And ξ_n converges in distribution towards a χ^2-distribution with $(r-1)(l-1)$ d.o.f.
The test is defined by the critical region:

$$W_n = \{\xi_n > q_{(r-1)(l-1)}(1-\alpha)\}$$

where $q_{(r-1)(l-1)}(1-\alpha)$ is the quantile of order $(1-\alpha)$ of the χ^2-distribution with $(r-1)(l-1))$ d.o.f. This test is strongly convergent at asymptotic level $\alpha = P(\chi^2((r-1)(l-1)) > q_{(r-1)(l-1)}(1-\alpha))$.

χ^2 test for independence

Example A study on 592 women: is there a correlation between eyes color and hairs color?

Eyes \ Hairs	Dark	Light-brown	Red	Blond
Black	68	119	26	7
Brown	15	54	14	10
Green	5	29	14	16
Blue	20	84	17	94

One obtains $\xi_n = 138{,}29$, $dof = 9$, $P(\chi_q^2 \leq 16{,}91) = 0{,}95$. Since $138{,}29 \gg 16{,}91$, one rejects H_0.