Maitrise statistique des procédés

2A AGB – Semestre 7 - 2019

Objectifs

Acquérir les méthodes et les outils d'application de la Maitrise Statistique des procédés (MSP).
À l'issue de cet enseignement, un étudiant sera capable de :
* présenter de manière synthétique des résultats à l'aide de paramètres et graphiques
* traiter des données, mettre en œuvre et interpréter des tests statistiques pour prendre des décisions
* mettre en place des outils de suivi de production.

Evaluation

S1 : 1 h, CR sur machine en utilisant R)

Contenu (4 cours, 6 TD)

Enseignant : Florent ARNAL

Adresse électronique : florent.arnal@u-bordeaux.fr

Site internet : http ://flarnal.e-monsite.com

PLAN DU COURS

Partie I : Tests statistiques
Tests de comparaisons : moyennes, variances et proportions
Tests non paramétriques
Analyse de variance : ANOVA à un facteur, deux facteurs
Applications : contrôle de préemballés, Analyse sensorielle,

Partie II : Maitrise Statistique des procédés
Généralités
Règle des 6 sigmas
Représentations graphiques (Diagrammes de causes et effets, diagramme de Pareto, …)
Capabilité d'une machine, d'un processus
Cartes de contrôle

1

PARTIE I : TESTS STATISTIQUES

I.1. Test du Khi-deux d'homogénéité & d'indépendance

Exemple :
Trois produits sont évalués par des consommateurs qui doivent répondre à la question suivante :
Achèteriez-vous ce produit ?
Les résultats obtenus sont les suivants :

	P1	P2	P3
OUI	40	50	60
NON	60	50	40

L'appréciation est-elle liée au produit ?

2

Déterminons les effectifs "théoriques" si l'appréciation est indépendante du produit.
On rappelle que si A et B sont indépendants alors $\mathbb{P}(A \cap B) = \mathbb{P}(A).\mathbb{P}(B)$.

Test du Khi-deux d'homogénéité & d'indépendance

$\mathcal{H}_0$: Les caractères étudiés sont indépendants
$\mathcal{H}_1$: Les caractères étudiés ne sont pas indépendants

On suppose que les deux caractères prennent respectivement p et q modalités.

Sous $\mathcal{H}_0$, la variable et la loi de décision sont les suivantes :

$$\sum_{i,j} \frac{\left(N_{ij} - N_{ij}^{th}\right)^2}{N_{ij}^{th}} \text{ suit la loi } \chi^2\left((p-1)\times(q-1)\right)$$

N_{ij} et N_{ij}^{th} correspondent respectivement aux effectifs observés et théoriques (cf. indépendance).
Ce test est par nature **unilatéral**.

Contrainte : les effectifs théoriques doivent être supérieurs à 5.
Dans le cas contraire, on doit effectuer des regroupements de classes ...

Mise en place du test

Avec R (Voir doc Begin'R)

```
Tableau = matrix(c(40, 50, 60, 60, 50, 40), nrow=2, ncol=3, byrow=TRUE)
colnames(Tableau) = c("P1", "P2", "P3")
rownames(Tableau) = c("OUI", "NON")
Tableau
chisq.test(Tableau)
```

```
Pearson's Chi-squared test

data:  Tableau
X-squared = 8, df = 2, p-value = 0.01832
```

5

I.2. Tests de comparaisons

On se propose de tester l'hypothèse selon laquelle dans deux populations P_1 et P_2 , il y a la même moyenne (μ inconnue) pour la variable étudiée à partir d'échantillons extraits de tailles respectives n_1 et n_2.

On a donc l'hypothèse nulle suivante :

$$\mathcal{H}_0 : \mu_1 = \mu_2$$

Plusieurs cas peuvent se présenter
(il faut également s'intéresser à la taille de l'échantillon) :

- les **variances** sont **connues**
- les **variances** sont **inconnues**
- les **échantillons** sont **appariés** (pas indépendants)

6

Tests de comparaison de 2 MOYENNES
Cas de distributions normales

- Cas où les échantillons sont **indépendants** et les **variances connues** :

$$\bar{X}_1 - \bar{X}_2 \sim \mathcal{N}\left(\mu_1 - \mu_2; \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$$

- Cas où les échantillons sont **indépendants** et les **variances inconnues mais** supposées **égales** *(hypothèse d'homoscédasticité)* :

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2} \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim \mathcal{T}\left(n_1 + n_2 - 2\right)$$

$\frac{SCE_1 + SCE_2}{n_1 + n_2 - 2} = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$ est une **estimation de la variance commune**

Avec R :
t.test(x= , y = , var.equal = TRUE, …)

7

Tests de comparaison de 2 MOYENNES

- Cas où les échantillons sont **indépendants** et les **variances inconnues mais différentes** :

Utilisation du test de Welch (adaptation d'un test de Student)
t.test(x= , y = , var.equal = FALSE, …)

- Cas où les échantillons ne sont **pas indépendants** (échantillons appariés) avec $n_1 = n_2 = n$:

On raisonne sur la distribution des différences entre les mesures et on utilise que :

$$\frac{\bar{X} - \mu}{\frac{\widehat{S}}{\sqrt{n}}} \hookrightarrow T\left(n - 1\right)$$

Mise en place du test avec R :
t.test(x= , y = , paired = TRUE, …)
Sous réserve de *normalité de la distribution des différences* !

8

Tests de comparaison de 2 PROPORTIONS

On se propose de tester l'hypothèse selon laquelle dans deux populations, il y a la même proportion (p inconnue) pour la variable mesurée.

On a donc l'hypothèse nulle suivante :
Ho : « $\mathbf{p_1 = p_2}$ »
Nous ne considérerons que le cas de **grands échantillons**.

Sous $\mathcal{H}_0$, on a

$$F_1 - F_2 \sim \mathcal{N}\left(0; \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\right)$$

où $\hat{p}$ est une estimation de la proportion commune.
Avec R : prop.test(x= , y= , ...)

9

Tests de comparaison de 2 VARIANCES

On se propose de tester l'hypothèse selon laquelle dans deux populations **gaussiennes indépendantes**, il y a la même variance pour la variable mesurée.

$$\mathcal{H}_0 : \sigma_1^2 = \sigma_2^2$$

On utilise un test de Fisher dont la statistique de test, sous $\mathcal{H}_0$ est

$$\frac{\hat{S_1}^2}{\hat{S_2}^2} \sim \mathcal{F}\left(n_1 - 1; n_2 - 1\right)$$

Mise en place du test avec R :
var.test(x= , y = , …)

10

I.3. Analyse de variance (ANOVA) et tests post hoc

L'ANOVA (Analysis of variance) :
Modèle linéaire gaussien dans lequel toutes les variables explicatives sont qualitatives.
Elles sont appelées **facteurs** (cf. plans factoriels) et leurs modalités sont appelées **niveaux** (contrôlés ou provoqués).
La variable aléatoire réponse est toujours **quantitative** et supposée **gaussienne**.

Exemple : (Effet d'une molécule activateur d'enzyme)
• 5 groupes de 10 souris.
• Doses : 1ng, 10 ng, 50 ng et 100 ng. Le 5e groupe ne reçoit que le solvant utilisé.

11

Pourquoi faire une ANOVA ?

- *Plusieurs tests réalisés* pour répondre à un problème donné (différents critères quantitatifs ou qualitatifs)

ou

- *Comparaison multiples de moyennes*

Exemple :
Comparaison de 4 traitements sans effet « réel »

Il y a 6 tests de comparaison à effectuer.
Le risque alpha global associé à ce test correspond à la probabilité de conclure qu'au moins un des traitements à un effet significatif.

12

Calcul du risque α_{global} associé à 6 tests :
$1-\alpha_{global}$ correspond à la probabilité qu'aucun des 6 tests ne mette en évidence un effet significatif (en l'absence d'effet).
Avec un seuil de signification de 5%, on a :

$$1 - \alpha_{global} = \cdots$$

$$\alpha_{global} = \cdots$$

De façon générale, pour n tests avec un seuil de signification α, on a :

$$\alpha_{global} = \cdots$$

Application pour la comparaison de 5 populations avec $\alpha = 5\%$: $\alpha_{global} \simeq 40\%$

Cas où α est voisin de 0 :

$$\alpha_{global} \sim \cdots$$

Conditions d'application

- Indépendances des p populations dont sont extraits les échantillons (cf. niveaux)
- **Homoscédasticité** et **Distributions gaussiennes** vérifiées sur les **résidus** usuellement

Formulation des hypothèses

$\mathcal{H}_0$: « $\mu_1 = \mu_2 = ... = \mu_p$ »

ie « il n'y a pas d'effet du facteur étudié »

$\mathcal{H}_1$: « Au moins deux moyennes sont différentes »

ie « il y a un effet du facteur étudié »

Principe de L'ANOVA

Modèle

p niveaux, n_i répétitions pour le niveau i, avec $N = \sum_{i=1}^{p} n_i$

$$
\left\{
\begin{array}{llll}
x_{11}\ x_{12} & \ldots & x_{1n_1} & \longrightarrow \ \bar{x}_1 \\
x_{21}\ x_{22} & \ldots & x_{2n_2} & \longrightarrow \ \bar{x}_2 \\
 & & & \\
x_{i1}\ x_{i2} & \ldots\ x_{ij} & x_{in_i} & \longrightarrow \ \bar{x}_i \\
 & & & \\
x_{p1}\ x_{p2} & \ldots & x_{pn_p} & \longrightarrow \ \bar{x}_p \\
\end{array}
\right.
$$

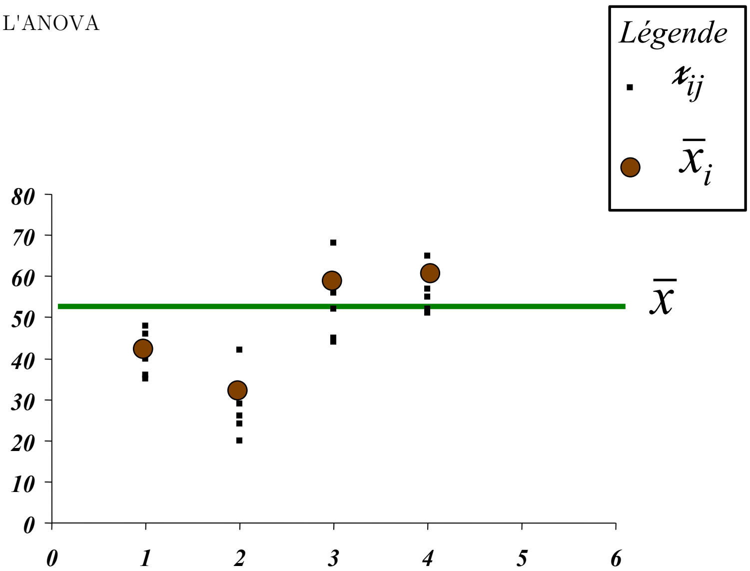
$j^{ème}$ *répétition du* $i^{ème}$ *traitement*

Moyenne du $i^{ème}$ *traitement*

Moyenne globale $\longrightarrow \bar{x}$

15

Principe de L'ANOVA



16

$$x_{ij} - \bar{x} = (x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{x}) \text{ avec } e_{ij} = x_{ij} - \bar{x}_i$$

donc $$x_{ij} = \bar{x} + (\bar{x}_i - \bar{x}) + e_{ij}$$

Observation

Part expliquée par l'effet du traitement

Moyenne globale

Résidu e_{ij}

Les résidus $e_{ij} = x_{ij} - \bar{x}_i$, par niveau (et donc au global), sont de somme et de moyenne nulle.

MODELE LINEAIRE

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

où

- μ est la moyenne (globale) associée aux p populations.
- X_{ij} est une variable aléatoire (réponse quantitative). Les X_{ij} sont indépendantes.
- ε_{ij} est une variable aléatoire d'erreur (non observées). On a : $\varepsilon_{ij} \sim \mathcal{N}(0; \sigma)$, σ étant un paramètre inconnu (à estimer).
- α_i correspond à l'effet associé à la modalité i avec $\sum_{i=1}^{p} \alpha_i = 0$.

  Une estimation de cet effet est donnée par

  $$\hat{\alpha}_i = \bar{x}_i - \bar{x}$$

Le modèle peut également s'écrire

$$X_{ij} = \mu_i + \varepsilon_{ij}$$

où $\mu_i = \mu + \alpha_i$ dont une estimation est $\hat{\mu}_i = \bar{x}_i$.

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

où

- X_{ij} est une variable aléatoire (réponse quantitative).
  Les X_{ij} sont indépendantes et gaussiennes.

  $$X_{ij} \sim \mathcal{N}(\mu_i; \sigma)$$

- ε_{ij} est une variable aléatoire d'erreur (non observées). On a :

  $$\varepsilon_{ij} \sim \mathcal{N}(0; \sigma)$$

  σ étant un paramètre inconnu avec

  $$\hat{\sigma}^2 = CM_{res}$$

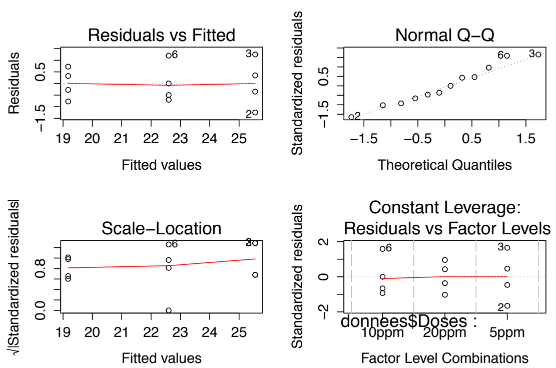
Les hypothèses de normalité et homoscédasticité se vérifient sur les résidus.
L'indépendance est liée au contexte de l'expérimentation.

19

Fonction plot sous R

Obtention de 4 graphiques permettant d'obtenir des informations sur :

✓ l'homoscédasticité (dispersion des résidus),
✓ la normalité (QQ-plot),
✓ l'indépendance des résidus.



Avec R :
anova = aov(var ~ facteur)
par(mfrow = c(2,2))
plot(anova)

20

Pour vérifier les conditions d'utilisation :

1. Égalité des variances :
   Méthodes graphiques, tests de Bartlett, Levene ou Brown-Forsythe avec

$$\mathcal{H}_0 : \sigma_1^2 = \cdots = \sigma_p^2$$

   - Test de Bartlett (couramment utilisé, notamment en cas de distributions gaussiennes) avec une statistique distribuée suivant une loi du Khi-deux.

   - Test de Levene correspondant à une anova sur les valeurs absolues des résidus.

   - Test de Brown-Forsythe, variante du test de Levene sur les médianes et non les moyennes.
     Gain de puissance en cas de non normalité.

2. Normalité de la distribution des résidus :
   Graphique Quantile-Quantile, test de Shapiro avec

$$\mathcal{H}_0 : \text{" La distribution est normale "}$$

21

En posant

- $SCE_{fact} = SCE_{inter} = \sum_{i,j} (\bar{x}_i - \bar{x})^2$
- $SCE_{res} = SCE_{intra} = \sum_{i,j} (x_{ij} - \bar{x}_i)^2$
- $SCE_{totale} = \sum_{i,j} (x_{ij} - \bar{x})^2$.

On a

$$SCE_{totale} = SCE_{fact} + SCE_{res}$$

Source	Df	SCE	CM	F_{obs}	p-valeur
Factorielle	$p-1$	SCE_{fact}	$CM_{fact} = \frac{SCE_{fact}}{p-1}$	f_{obs}	$\mathbb{P}(F_{obs} > f_{obs})$
Résiduelle	$N-p$	SCE_{res}	$CM_{res} = \frac{SCE_{res}}{N-p}$		
Totale	$N-1$	SCE_{totale}			

pour p modalités, N observations. Sous $\mathcal{H}_0$,

$$F_{obs} = \frac{CM_{fact}}{CM_{res}} \sim \mathcal{F}(p-1; N-p)$$

22

Si f_0 **est proche de 0** :
Variabilité factorielle inférieure à la variabilité résiduelle.
Pas d'effet factoriel significatif puisque la variabilité totale est due essentiellement à la variabilité résiduelle.

Ainsi : par nature, ce test sera toujours **unilatéral**.

Si on ne rejette pas $\mathcal{H}o$ alors on peut dire que le facteur étudié n'a pas plus d'effet que les variations aléatoires.

Si on rejette $\mathcal{H}o$ alors on peut dire que le facteur étudié a un effet significatif. Il est alors intéressant de comparer ces différentes moyennes à l'aide, par exemple, de la ppds.

23

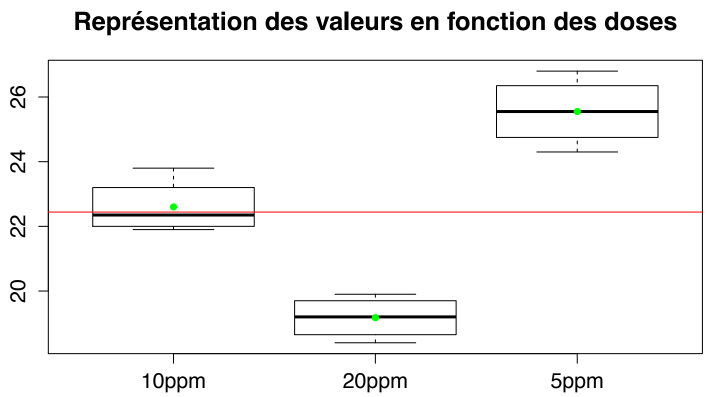
Exemple

Exemple :
L'acide citrique a-t-il une influence sur le brunissement de pâtes alimentaires ?

	Doses d'acide citrique		
	5 ppm	**10 ppm**	**20 ppm**
Indices de brun	**25,2**	**22,1**	**18,4**
	24,3	**23,8**	**19,5**
	26,8	**21,9**	**18,9**
	25,9	**22,6**	**19,9**

Facteur étudié : Quantité d'acide citrique
Nombre de niveaux : 3
Plan équilibré (même nombre de répétitions)

24



ANALYSE DE
VARIANCE

Source des variations	*Somme des carrés*	*Degré de liberté*	*Moyenne des carrés*	*F*	*Prob*
Entre Groupes	81,43				
A l'intérieur des groupes	6,86				
Total	88,29				

Avec R :
anova = aov(var ~ facteur)
summary(anova)

Comparaison de moyennes 2 à 2
Tests post-hoc

Test de Bonferroni (cf. test de Student) :

Présentation dans le cas d'échantillons de même taille n

Le carré moyen résiduel CM_{res} est un bon estimateur de la variance commune.

Sous l'hypothèse d'égalité des moyennes :

27

Pour chaque comparaison, on effectue un test avec un seuil de signification égal à

$$\frac{\alpha}{\text{nombre de tests}} = \frac{\alpha}{\binom{p}{2}}$$

où $\alpha = 0,05$ (usuellement) et $\binom{p}{2} = \frac{p(p-1)}{2}$.

Avec R :
pairwise.t.test(variable, facteur, p.adj = "bonferroni")
LSD.test( ) via le package agricolae

Un test plus puissant (surtout si plan équilibré) basé sur les plus petites amplitudes significatives(ppas) :
Test de Student-Newman-Keuls
Avec R :
SNK.test( ) via le package agricolae

28

I.4. Tests non paramétriques

- À utiliser lorsque conditions d'applications de certains tests ne sont pas vérifiées.
- Basés sur des signes ou des rangs.
- Moins puissants que des tests paramétriques.

Pour comparer deux moyennes :
Test des signes
Tests des rangs de Wilcoxon
wilcox.test( )

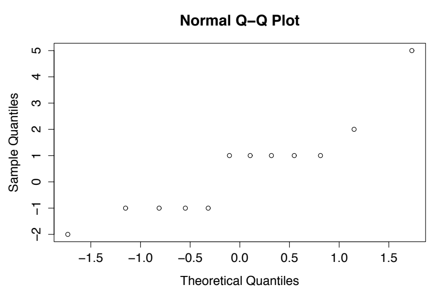
Pour comparer plus de 2 moyennes :
Test des rangs de Kruskal-Wallis
kruskal.test( )

29

Exemple avec le test des signes

Notes attribuées par 12 personnes d'un jury sur 2 produits

Produit 1	6	6	7	6	7	8	5	6	8	8	8	8
Produit 2	5	7	6	5	8	9	0	7	6	7	10	7



30

Exemple avec le test des signes

On va considérer les différences entres les notes des deux produits.
$\mathcal{H}_0$: "les notes sont similaires"

Sous $\mathcal{H}_0$, le nombre X de différences positives (ou négatives) est tel que

$$X \sim \mathcal{B}(n; 0,5)$$

Le test des signes se ramène à un test binomial.

Résultats avec R

➢ binom.test(x= …. , n= …. , p= …. , alternative= "greater")
Exact binomial test
data: 7 and 12
number of successes = 2, number of trials = 12,
p-value = 0.381

➢ **Avec le test de Wilcoxon**
wilcox.test(x, y, paired=……….. , alternative = "greater")
data: x and y
p-value = 0.255

PARTIE II : MSP

II.1. Généralités

- **Objectif général**
  Anticiper sur les mesures à prendre pour améliorer n'importe quel processus de fabrication industrielle

- **Historique**
  Mise en place au Japon après la Seconde Guerre mondiale grâce à Deming, disciple de Shewhart.
  Développée aux USA dans les années 60, en Europe dans les années 80

- **Outils**
  Contrôle de réception, les plans d'expérience, la capabilité, les cartes de contrôle, etc.

33

➡ Mise en place de contrôles en cours de production

Objectif :

Obtenir une production stable avec un minimum de produits non conformes aux spécifications.

➡ Contrôle de la qualité « dynamique »

il ne s'intéresse pas au résultat isolé et instantané, mais au suivi dans le temps : il ne suffit pas qu'une pièce soit dans les limites des spécifications, il faut aussi surveiller la répartition chronologique des pièces à l'intérieur des intervalles de tolérances (cf.dérives).

La MSP (Statistical Process Control ou SPC) a pour objet une qualité accrue par l'utilisation d'outils statistiques visant à une production centrée et la moins dispersée possible.

Objectifs de la MSP

- Améliorer la confiance du client
- Diminuer les coûts opératoires
- Réduire les contrôles
- Améliorer le procédé
  - Diminuer la dispersion
  - Détecter les dérives

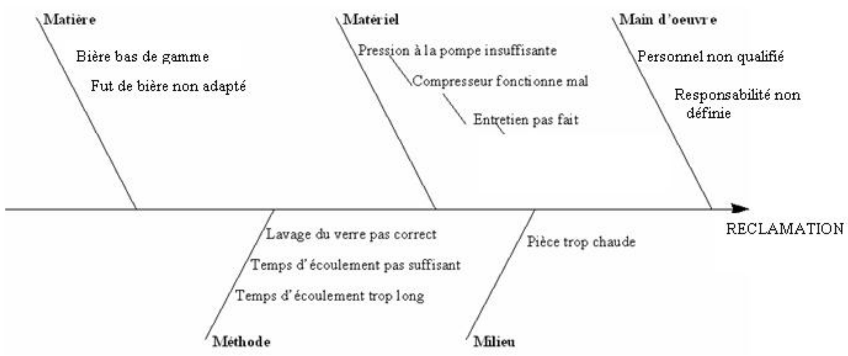
Les 5M : Causes fondamentales de la dispersion (Ishikawa)

- Matière(s)
- Machine(s)
- Main d'œuvre
- Méthode(s)
- Milieu

On parle parfois des 6M en considérant la Mesure
(ne modifie pas la dispersion de la production vendue au client mais l'image que l 'on en a)

Les 5M : Causes fondamentales de la dispersion (Ishikawa)

Diagramme de Causes/effets :



Un exemple de problématique

Le cahier des charges d'une production de fromages de chèvre indique :
Masse : 1 kg $\pm$ 10 g ;
Longueur : 26 cm $\pm$ 0,5 cm.

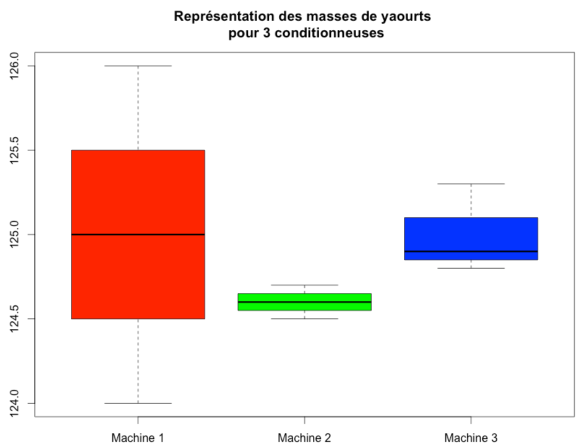
Sur un échantillon de 5 fromages, on a relevé les masses suivantes (en g) :

992 995 1000 1005 1008

Faut-il "s'inquiéter" de ses résultats ?
Quels seraient vos "objectifs" pour répondre aux exigences du cahier des charges concernant la longueur des bûches ?

Représentations graphiques « classiques »

Données quantitatives : le boxplot



Représentations graphiques « classiques »

Données qualitatives (Répartition des défauts d'une production) :
Le diagramme de Pareto

Ce diagramme, basé sur des fréquences cumulées, permet d'évaluer l'importance de différents facteurs sur un processus.

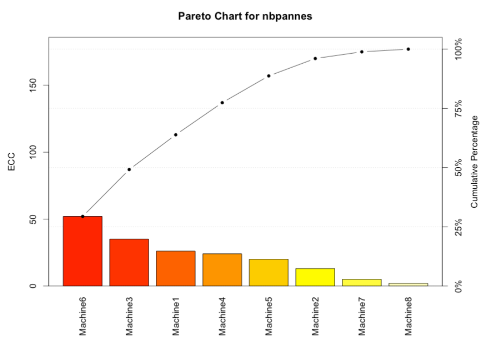
Historique : V. Pareto, économiste italien, avait fait une étude sur la répartition des richesses mettant en évidence que 80 % des richesses étaient détenues par 20 % de la population (loi des 80/20).

Juran en tire l'idée que, pour un phénomène, 20 % des causes produisent 80 % des effets (d'où l'intérêt de s'intéresser à la répartition des défauts d'une production).

Représentations graphiques « classiques »

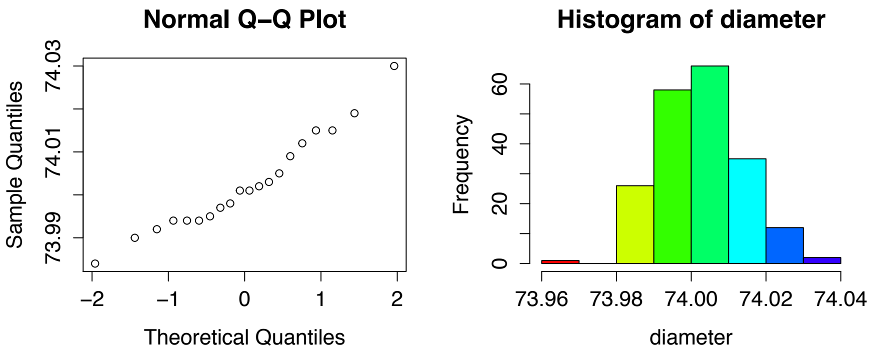
Données qualitatives (Répartition des défauts d'une production) :
Le diagramme de Pareto

Numéro De machine	Nombre de pannes sur 1 an
1	26
2	13
3	35
4	24
5	20
6	52
7	5
8	2



Avec R : *pareto.chart*( ) via le package qcc

Autour de la normalité d'une distribution



Dans ce qui suit, on considère des distributions normales

II.2. Règle des 6 sigma et Capabilité

Méthode 6 Sigma : Quel intérêt ?

Considérons une fabrication de pièce avec une longueur cible égale à L.
La pièce est utilisable si sa longueur appartient à $[L - \Delta L; L + \Delta L]$.

Si on considère que : $\Delta L = 3\sigma$ alors

$$P\left(X \in [L - \Delta L; L + \Delta L]\right) \simeq 0,9973$$

soit 2700 pièces défectueuses par million.

Si on considère que : $\Delta L = 6\sigma$ alors

$$P\left(X \notin [L - \Delta L; L + \Delta L]\right) \simeq 2 \times 10^{-9}$$

soit 2 pièces défectueuses par milliard.

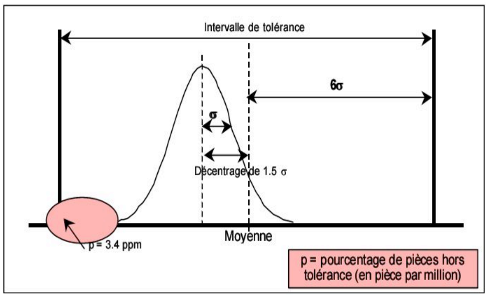
Comment atteindre cet objectif en production ?

- Augmenter les limites de tolérance
- Diminuer la variabilité de la production

Règle des 6 Sigma et Capabilité

Règle 6-sigma et dérive

L'approche Six Sigma tient compte d'une éventuelle déviation de $1,5\sigma$



$$P\left(X \notin [L - 4,5\sigma; L + 7,5\sigma]\right) \simeq 3,4 \times 10^{-6}$$

soit 3,4 pièces défectueuses par million.

Règle des 6 Sigma et Capabilité

Notion de capabilité

Rapport entre la performance réelle d'un procédé et la performance attendue (exigée)

La capabilité d'un processus associé à un intervalle de tolérance $[T_I; T_S]$ est définie par :

$$C_p = \frac{T_S - T_I}{6\sigma}$$

Dans le cas d'un processus avec une dérive sur la valeur cible, on introduit :

$$C_{pk} = \min\left\{\frac{T_S - \mu}{3\sigma}; \frac{\mu - T_I}{3\sigma}\right\}$$

On considère une machine "capable" lorsque ces coefficients sont supérieurs à **1,33** ...

Règle des 6 Sigma et Capabilité

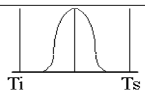
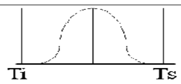
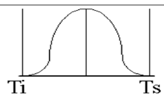
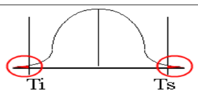
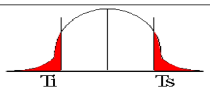
En phase de pré-industrialisation, ces indices font référence à la capabilité Machine. On les note alors :

$$C_m \quad \text{et} \quad C_{mk}$$

En phase de production, les indices de capabilité ne reflètent la qualité du procédé que lorsque celui-ci est sous contrôle (stable en moyenne et dispersion).

Fonction sous R :
process.capability( )
via le package qcc

Règle des 6 Sigma et Capabilité

1	Cp > 1.67		Plus que suffisant	Non préoccupant, chercher à simplifier la gestion pour réduire les coûts.
2	1.67 > Cp > 1.33		suffisant	Situation idéale. A maintenir.
3	1.33 > Cp > 1.00		trop juste	Nécessite de l'attention, Cp proche de 1 signifie qu'une dérive peut créer des défauts.
4	1.00 > Cp > 0.67		insuffisant	Existence de Non Conformes. il faut contrôler à 100%, analyser le processus et si possible l'améliorer.
5	0.67 > Cp		très insuffisant	Analyse immédiate des causes, urgence de mise en place de contre-mesures, révision des tolérance.

II.3. Cartes de contrôle

Présentation

Graphique sur lequel on reporte, dans l'ordre chronologique, les valeurs d'une statistique calculée sur des échantillons, en général de même effectif, issus de la fabrication.

Une telle carte comporte une ligne centrale (valeur cible) ainsi que des limites de surveillance et de contrôle tracées à l'avance.

- Abscisse : Numéro de l'échantillon
- Ordonnée : Valeur de la statistique calculée sur cet échantillon

Cartes de contrôle

Principe

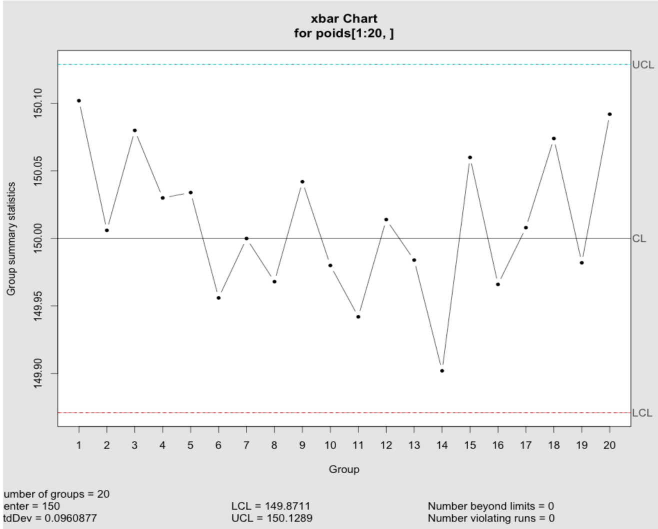
Graphique sur lequel on reporte, dans l'ordre chronologique, les valeurs d'une statistique calculée sur des échantillons, en général de même effectif, issus de la fabrication.

Une telle carte comporte :
- une ligne centrale (valeur cible)
- des limites de surveillance et de contrôle

Les points placés ont pour :
- Abscisse : Numéro de l'échantillon
- Ordonnée : Valeur de la statistique calculée sur cet échantillon

Cartes de contrôle

Carte de contrôle de Shewart sur la moyenne



Cartes de contrôle

Principe pour la carte de contrôle de Shewart de la moyenne

La valeur cible correspond à la vraie valeur du paramètre (ou son estimation ponctuelle).

Les limites de contrôle inférieures (LIC ou LCL en anglais) et supérieures (LSC ou UCL en anglais) sont situées à 3 écarts-types de la cible$\mathrm{LCL} = \mathrm{LIC} = \mu - 3\frac{\sigma}{\sqrt{n}}$; $\mathrm{UCL} = \mathrm{LSC} = \mu + 3\frac{\sigma}{\sqrt{n}}$.

Étant donné que $\overline{X}$ est distribuée suivant la loi normale $\mathcal{N}\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$, on a

$$P\left(\mu - 3\frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + 3\frac{\sigma}{\sqrt{n}}\right) \simeq 0,9973$$

Ainsi : la plupart des moyennes observées sont comprises entre ces 2 limites.

Cartes de contrôle

Principe pour la carte de contrôle de Shewart de la moyenne

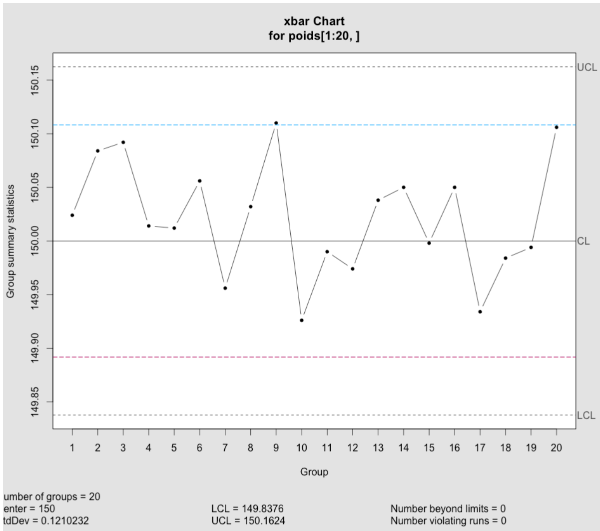
On peut rajouter des limites de surveillance situées à 2 écarts-types de la cible.

Étant donné que $\overline{X}$ est distribuée suivant la loi normale $\mathcal{N}\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$, on a

$$P\left(\mu - 2\frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + 2\frac{\sigma}{\sqrt{n}}\right) \simeq 0,954$$

Une observation comprise entre les limites de surveillance et de contrôle (inf ou sup) doit inciter à une attention particulière (prélèvement d'un nouvel échantillon par exemple).

Cartes de contrôle



Carte de contrôle de la moyenne avec limites de contrôle et surveillance

Avec R :

Package qcc

qcc(valeurs, type="xbar")

Cartes de contrôle

Étude et utilisation de la carte de contrôle de Shewart de la moyenne
Norme NF 06-031-1

1. Vérifier que la distribution des valeurs est normale.

2. Estimer l'écart-type de la production en utilisant les variances ou étendues (voir page suivante).

3. Calculer des limites de contrôle (et surveillance).

4. Suivi de la production
   On peut considérer qu'il y a un déréglage lorsque :
   - Point à l'extérieur des limites de contrôle ;
   - 9 points consécutifs du même côté de la ligne centrale ;
   - 6 points consécutifs tous « ascendants » ou « descendants ».

Cartes de contrôle

Estimation de l'écart-type par les étendues

1. Prélever plusieurs échantillons de même taille n (issus d'une population gaussienne).

2. Calculer pour chaque échantillon l'étendue R_i ainsi que l'étendue moyenne $\bar{R}$.

3. Déterminer d_2 en fonction de n.

   On a alors

$$\widehat{\sigma} = \frac{\bar{R}}{d_2}$$

Table 1: Control Chart Coefficients

Subgroup Size n	d_2	D_1	D_2	D_3	D_4	A_2
2	1.128	0	3.686	0	3.267	1.880
3	1.693	0	4.358	0	2.575	1.023
4	2.059	0	4.698	0	2.282	0.729
5	2.326	0	4.918	0	2.115	0.577
6	2.534	0	5.078	0	2.004	0.483
7	2.704	0.205	5.203	0.076	1.924	0.419
8	2.847	0.387	5.307	0.136	1.864	0.373
9	2.970	0.546	5.394	0.184	1.816	0.337
10	3.078	0.687	5.469	0.223	1.777	0.308
11	3.173	0.812	5.534	0.256	1.744	0.285
12	3.258	0.924	5.592	0.284	1.716	0.266
13	3.336	1.026	5.646	0.308	1.692	0.249
14	3.407	1.121	5.693	0.329	1.671	0.235
15	3.472	1.207	5.737	0.348	1.652	0.223
20	3.735	1.548	5.922	0.414	1.586	0.180
25	3.931	1.804	6.058	0.459	1.541	0.153

Cartes de contrôle

Pour compléter l'étude sur les moyennes : Étude de la variabilité

1. Carte de contrôle de l'étendue (qcc(valeurs, type= "R"))
   Carte $(\bar{x}, R)$

1. Carte de contrôle de l'écart-type (qcc(valeurs, type= "S"))
   Carte $(\bar{x}, s)$

Pour contrôler des proportions ou nombre de défectueux (qualitatif) :
Carte de contrôle aux attributs

1. Carte de contrôle de la proportion (p)
   qcc(valeurs, type= "p")

2. Carte de contrôle du nombre de défectueux (np)
   qcc(valeurs, type= "np")

3. Carte de contrôle du nombre de défauts (c)
   qcc(valeurs, type= "c"))

Application 1

Déterminer les limites de contrôle d'une carte sur la proportion p ?

Application 2

On a prélevé 5 échantillons de 4 fromages.

1. Déterminer, par calcul, une estimation de l'écart-type de la production en utilisant les étendues.
2. Déterminer les limites de contrôle de la carte de Shewart des moyennes.
3. Déterminer les limites de surveillance.
4. L'entreprise a fixé une valeur nominale de 100 g avec des tolérances de 99 à 101 g.
   Déterminer les indices de capabilité C_p et C_{pk}.

Numéro	Moyenne	Étendue
1	100	0,6
2	100,25	0,4
3	99,55	0,4
4	99,65	0,5
5	100,4	0,6

Application 3

On s'intéresse à une production de bouteilles de jus de fruits (QN = 150 cl) avec un tiers de Mangue, un tiers d'Orange et un tiers d'eau.
On suppose que les distribution des quantités de chaque constituant sont normales, d'espérance 50 et d'écart-type σ. Le cahier des charges impose une production contenant 150 $\pm$ 0,5 cl.

1. Dans cette partie, on suppose que $\sigma = 0.2$ cl.
   (a) Créer des échantillons de quantités associés à la mangue, l'orange et l'eau. Proposer une représentation graphique des quantités associées à ces trois constituants.
   (b) En déduire un échantillon de volumes associés à 50 bouteilles de cette production.
   (c) La distribution des quantités des bouteilles de la production est-elle normale ?
   (d) Déterminer l'écart-type des volumes des bouteilles de cette production.
   (e) La production sera-t-elle jugée conforme ?
2. Que préconiseriez-vous ?