

Douceurs statistiques

Comment rendre les statistiques plus appétissantes...

Rares sont ceux qui n'aiment pas les M&M's[®]. Plus nombreux sont ceux qui n'aiment pas les statistiques. Cet article est une recette pour rendre les statistiques plus « appétissantes », grâce aux M&M's !

En pesant quelques sachets de M&M's et en comptant leur contenu, vous collecterez des données très utiles. Les tests statistiques sur ces données vous permettront de répondre aux questions sur le poids, la couleur, la qualité et la quantité. Évidemment, si vous vous délectez à l'idée de croquer les confiseries à l'issue des tests, vous serez très motivés par cet exercice...

1^{ère} question : le poids indiqué sur le sachet de M&M's est-il exact ?

Chaque sachet de cacahuètes enrobées de chocolat devrait peser 49,3 grammes. Pour déterminer si le poids des sachets est conforme au poids cible, pesez 30 sachets de M&M's et effectuez un test t à un échantillon.

Test T à un échantillon : Poids

Test de mu = 49,3 en fonction de la différence 49,3

Variable	N	Moyenne	EcTyp	ErT moyenne	IC à 95 %	T	P
Poids	30	52,040	2,807	0,512	(50,992; 53,088)	5,35	0,000

Résultats obtenus dans le logiciel de statistiques Minitab® 15 en français à partir de la séquence de commandes **Stat > Statistique élémentaires > Test t à 1 échantillon**.

Les résultats du test t incluent une valeur de p de 0,000. La valeur de p étant inférieure à la valeur de α de 0,05, vous rejetez l'hypothèse nulle et concluez que le poids moyen de chaque sachet n'est *pas* 49,3 g. Vous constatez un poids moyen de 52,040 g, *supérieur* au poids cible de 49,3 grammes, c'est une bonne nouvelle pour tous les gourmands.

2^{ème} question : si une autre personne pèse les sachets, obtient-elle les mêmes résultats ?

Si deux personnes pèsent à tour de rôle 30 sachets de M&M's à l'aide de la même balance, obtiennent-elles des poids similaires ? Un test t apparié vous aide à répondre.

Test T pour données appariées et IC : Poids Personne1; Poids Personne2

Test t pour données appariées pour Poids Personne1 - Poids Personne2

	N	Moyenne	EcTyp	ErT moyenne
Poids Personne1	30	52,040	2,807	0,512
Poids Personne2	30	52,047	2,802	0,512
Différence	30	-0,00700	0,04435	0,00810

Différence des moyennes : limites de confiance (à 95 %) : (-0,02356 ; 0,00956)

Test t de la différence des moyennes = 0 (en fonction de la différence 0) : T = -0,86 P = 0,394

Résultats obtenus à partir de la séquence de commandes **Stat > Statistiques élémentaires > Test t pour données appariées.**

La valeur de p de 0,394 est supérieure à la valeur de α de 0,05. L'hypothèse nulle ne peut être rejetée : il n'existe pas de différence significative entre les mesures moyennes prises par les deux personnes.

3^{ème} question : le logo « m » est-il imprimé sur tous les M&M's ?

Les M&M's s'identifient facilement par leur marque, un petit « m » imprimé sur chaque bonbon. Imprimer correctement ce logo sur une surface aussi inégale que celle d'une cacahuète enrobée de chocolat est une prouesse technique.

Vous supposez, par exemple, que le taux de M&M's dont le logo est mal imprimé est égal ou inférieur à 15%. Pour vérifier cette hypothèse, comptez, d'une part le nombre total de M&M's, d'autre part ceux dont la marque « m » n'est pas imprimée correctement; puis effectuez un test à une proportion dans le logiciel de statistiques Minitab.

Test et IC pour une proportion

Test de $p = 0,15$ contre valeur de p supérieure à $0,15$

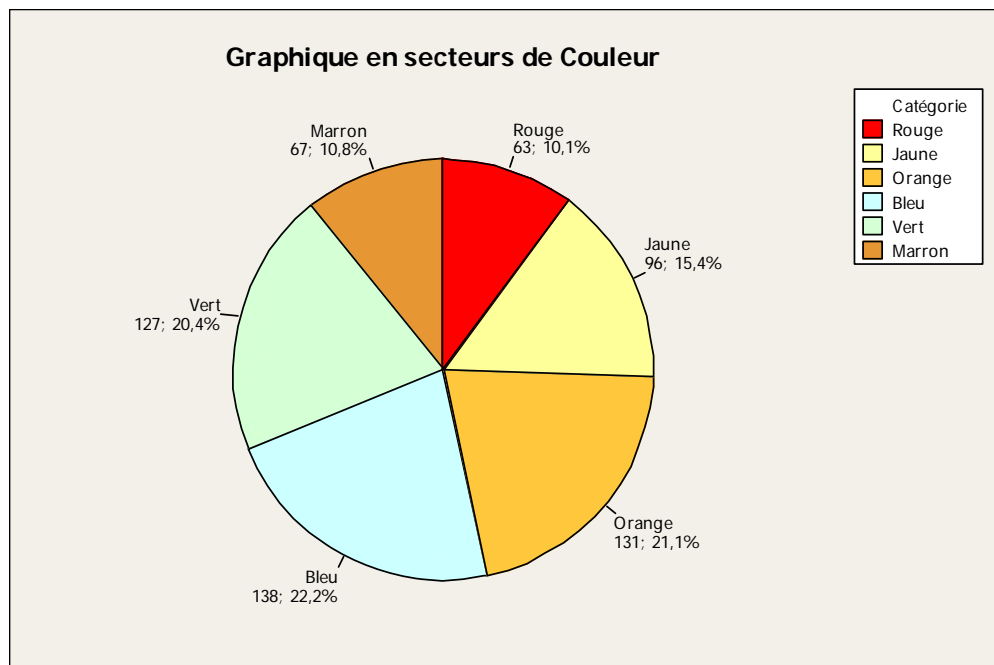
Echantillon	X	N	P échantillon	Borne inférieure = 95 %	Valeur exacte de P
1	87	622	0,139871	0,117497	0,776

Résultats obtenus à partir de la séquence de commandes **Stat > Statistiques élémentaires > 1 proportion.**

87 des 622 M&M's présentent un tel défaut. En utilisant un test pour une proportion et une hypothèse alternative de plus 15 %, vous obtenez une valeur de p de 0,776. La valeur de p étant supérieure à un seuil de alpha α égal à 0,05, la proportion de chocolats dont le logo présente un défaut, est ainsi inférieure ou égale à 15%.

4^{ème} question : un sachet contient-il un nombre égal de couleurs ?

Un graphique en secteurs permet de visualiser le nombre de M&M's de chaque couleur.



Commandes **Graphique > Graphique en secteurs** dans le logiciel Minitab.

L'échantillon contient 138 M&M's bleus et 63 M&M's rouges. Cette différence est-elle significative statistiquement ? Utilisez le test d'ajustement du Khi deux pour le savoir.

Test d'ajustement du Khi deux pour les dénombrements observés dans la variable : Nombre par couleur

Utilisation des noms de catégories dans Couleur

Catégorie	Observé	Proportion testée	Attendu	Contribution au Khi deux
Rouge	63	0,166667	103,667	15,9528
Jaune	96	0,166667	103,667	0,5670
Orange	131	0,166667	103,667	7,2069
Bleu	138	0,166667	103,667	11,3708
Vert	127	0,166667	103,667	5,2519
Marron	67	0,166667	103,667	12,9689

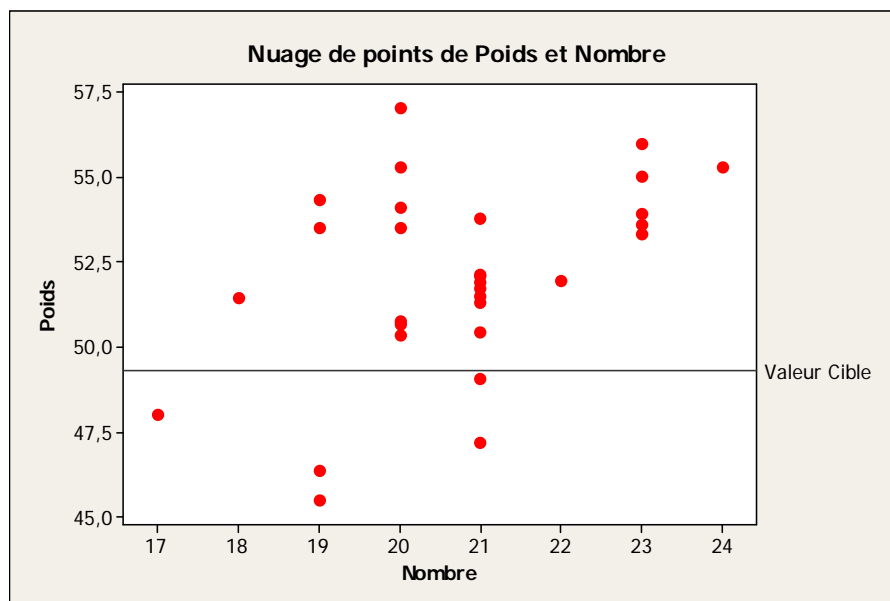
N	DL	Khi deux	Valeur de p
622	5	53,3183	0,000

Résultats obtenus à l'aide de la séquence de commandes **Stat > Tableaux > Test d'ajustement du khi deux (Une Variable)**.

La valeur de p de 0,000 confirme la première impression : les nombres de M&M's rouges, oranges, jaunes, verts, bleus et marrons ne sont pas égaux.

5^{ème} question : existe-t-il une corrélation entre le nombre de M&M's dans chaque sachet et le poids du sachet ?

Lorsque le nombre de M&M's augmente, le sachet s'alourdit-il ? Un diagramme à nuage de points et un test de corrélation révéleront la relation entre quantité et poids.



Commandes dans le logiciel Minitab : **Graphique > Nuage de points**

Corrélations : Poids; Nombre

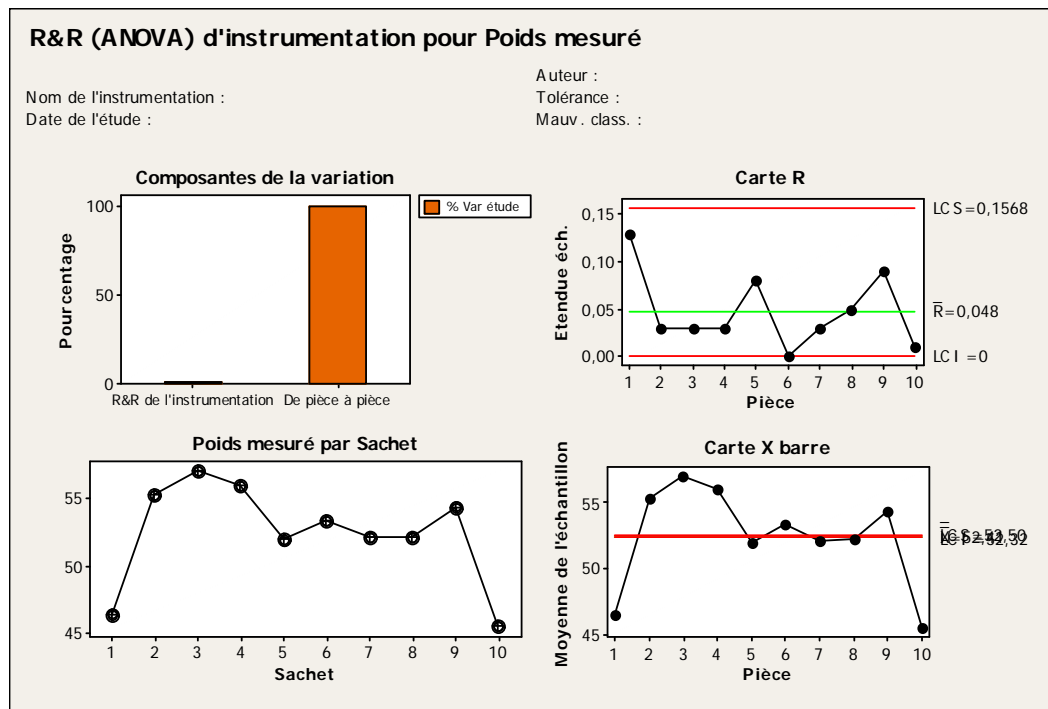
Corrélation de Pearson de Poids et Nombre = 0,458
 Valeur de p = 0,011

Résultats obtenus à partir de la séquence de commandes **Stat > Statistiques élémentaires > Corrélation.**

Selon la représentation en nuage de points des données, le poids augmente si la quantité de M&M's augmente. En outre, la valeur de p de 0,011 indique que l'hypothèse nulle peut être rejetée : il existe une corrélation positive entre le poids des sachets et le nombre de M&M's qu'ils contiennent.

Considérations supplémentaires

Vous pouvez également utiliser les M&M's pour enseigner ou apprendre les techniques statistiques plus avancées telles que l'analyse du système de mesure à l'aide de l'étude de la répétabilité et de la reproductibilité. Avec le logiciel Minitab, créez un plan de collecte de données pour l'étude R&R de l'instrumentation. Ce plan vous guide dans la mesure de 10 sachets de M&M's, chacun à deux reprises, dans un ordre aléatoire. Puis, les résultats obtenus sont analysés dans le logiciel :



R&R de l'instrumentation

Source	Ecart type	Variations	
		d'étude (6 * écart type)	% var étude (%VE)
R&R d'instrumentation total	0,04336	0,2602	1,14
Répétabilité	0,04336	0,2602	1,14
De pièce à pièce	3,81028	22,8617	99,99
Variation totale	3,81052	22,8631	100,00

Nombre de catégories distinctes = 123

Résultats obtenus à l'aide de la séquence de commandes **Stat >Outils de la Qualité > Étude de l'Instrumentation > Étude R&R de l'instrumentation (Croisée).**

La variation est évaluée à 1,14 %, soit un pourcentage très inférieur à l'idéal de 10 % recherché dans l'étude R&R. Une très faible variation est donc due au système de mesure. Tous les points de la carte R se trouvent dans les limites de contrôle. Ces deux éléments confirment la fiabilité du système de mesure.

Conclusion

Nous espérons que vous apprécierez cet exemple (et les M&M's) autant que nous. Cet exercice peut illustrer vos actions de formation ou vos cours, la leçon sera délicieuse !

Michelle Paret

Product Manager, Minitab Inc.

Eston Martz

Senior Creative Services Specialist, Minitab Inc.



M&M's[®] est une marque déposée de Mars, Incorporated Family of Companies

Visitez le site Web www.minitab.fr pour de plus amples informations sur la statistique ou contactez-nous par téléphone au + 33 (0)1 55 33 12 36 ou par courrier électronique à l'adresse bienvenue@minitab.fr.