Calcul de probabilité d'une loi normale de moyenne inconnue.

Enoncé du problème (problème inventé)

Un industriel remplit en série de bouteilles d'eau d'un litre à partir d'un appareil de remplissage dont l'écart type est de 25 ml.

Le produit mis en vente doit absolument contenir 1000 ml d'eau.

L'organisme national de contrôle peut procéder à sa convenance à des tests de conformité sur le volume d'eau contenue dans chaque bouteille.

Pour ce faire, ce même organisme a défini un test de conformité qui est le suivant:

1. Une bouteille d'eau est prise au hasard dans la production; si son volume d'eau est supérieur ou égal à 1000 ml, alors le test est réussi, sinon
2. 3 bouteilles d'eau sont prises au hasard dans la production; si la moyenne d'eau des 3 bouteilles est supérieure ou égale à 1000 ml, alors le test est réussi sinon le test de conformité échoue et toutes les bouteilles d'eau doivent être retirées du marché.

Afin de se prémunir d'un tel risque, l'industriel demande à un de ses ingénieurs de calculer le risque qu'il prend en réglant l'appareil de remplissage de telle sorte que les bouteilles contiennent 1020 ml d'eau.

Résolution du problème (solution sans incertitude sur la moyenne)

La probabilité P que le test de conformité dans sa globalité échoue est égal à P1T * P3T où

$P_{1T} = \int_{y=-\infty}^{LIMIT} N(y, \mu, \sigma) dy$ = probabilité que le premier test échoue

$P_{3T} = \int_{z=-\infty}^{LIMIT} N\left(z, \mu, \frac{\sigma}{\sqrt{3}}\right) dz$ = probabilité que le second test échoue (moyenne 3 mesures)

où

$$N(x, \mu, \sigma) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

μ = moyenne de la loi normale
σ = écart type de la loi normale
N = fonction de densité de probabilité de la loi normale
LIMIT = limite à ne pas dépasser = 1000 ml

L'équation de la probabilité peut alors s'écrire sous la forme suivante

$$P = \int_{y=-\infty}^{LIMIT} N(y, x, \sigma) dy \int_{z=-\infty}^{LIMIT} N(z, x, \frac{\sigma}{\sqrt{3}}) dz$$

Résolution du problème (solution 1)

Le problème est que la moyenne (1020 ml) est une moyenne issue de N mesures (N est petit, 1 ou 2 ou 3 mesures) et qu'il existe une certaine incertitude sur la mesure.

Si l'on tient compte du fait que la moyenne est incertaine et qu'elle varie selon une même loi normale, l'équation de la probabilité P_{1T} peut s'écrire sous la forme suivante:

$$P_{1T} = \int_{x=-\infty}^{+\infty} \left(N\left(x, \mu, \frac{\sigma}{\sqrt{n}}\right) \int_{y=-\infty}^{LIMIT} N(y, x, \sigma) dy \right) dx$$

Si l'on applique cette façon de faire à l'équation P1T*P3T, on obtient alors la probabilité suivante:

$$P_1 = \int_{x=-\infty}^{+\infty} \left(N\left(x, \mu, \frac{\sigma}{\sqrt{n}}\right) \int_{y=-\infty}^{LIMIT} N(y, x, \sigma) dy \int_{z=-\infty}^{LIMIT} N(z, x, \frac{\sigma}{\sqrt{3}}) dz \right) dx$$

Résolution du problème (solution 2)

Une autre façon de faire est de considérer que les 2 probabilités sont bien indépendantes et de calculer P_{1T} et P_{3T} séparément.

Dans ce cas, les formules des probabilités P1T et P3T sont les suivantes :

$$P_{1T} = \int_{x=-\infty}^{+\infty} \left(N\left(x, \mu, \frac{\sigma}{\sqrt{n}}\right) \int_{y=-\infty}^{LIMIT} N(y, x, \sigma) dy \right) dx$$

$$P_{3T} = \int_{x=-\infty}^{+\infty} \left(N\left(x, \mu, \frac{\sigma}{\sqrt{n}}\right) \int_{y=-\infty}^{LIMIT} N\left(y, x, \frac{\sigma}{\sqrt{3}}\right) dy \right) dx$$

Et la probabilité globale est:

$$P_2 = \int_{x=-\infty}^{+\infty} \left(N\left(x, \mu, \frac{\sigma}{\sqrt{n}}\right) \int_{y=-\infty}^{LIMIT} N(y, x, \sigma) dy \right) dx * \int_{x=-\infty}^{+\infty} \left(N\left(x, \mu, \frac{\sigma}{\sqrt{n}}\right) \int_{y=-\infty}^{LIMIT} N\left(y, x, \frac{\sigma}{\sqrt{3}}\right) dy \right) dx$$

Conclusion

La probabilité P_1 est différente de la probabilité P_2 !

Quelle est la meilleure estimation du risque ?

Pourquoi ?

Remarque

La formule de probabilité

$$P_{1T} = \int_{x=-\infty}^{+\infty} \left(N\left(x, \mu, \frac{\sigma}{\sqrt{n}}\right) \int_{y=-\infty}^{LIMIT} N(y, x, \sigma) dy \right) dx$$

peut être calculée en utilisant une autre loi normale dont la variance peut être déduite de l'équation suivante :

$${\sigma_{1T}}^2 = \sigma^2 + \frac{\sigma^2}{n}$$

Cela donne

$$\sigma_{1T} = \sigma \sqrt{\frac{n+1}{n}}$$

L'équation P_{1T} peut aussi être représentée par la formule suivante :

$$P_{1T} = \int_{y=-\infty}^{LIMIT} N\left(y, \mu, \sigma \sqrt{\frac{n+1}{n}} \right) dy$$