

Vous venez d'obtenir un résultat vous indiquant que :

**Résultat :**

Votre taux de défaillance prévisionnel donné par Fides  
n'est pas compatible avec votre retour d'expérience

[Plus de détails](#)

Nous allons dans ce complément d'information, vous indiquer la raison d'un tel résultat.

Dans ce test, nous cherchons à montrer si le taux de défaillance prévisionnel FIDES est compatible ou non avec notre retour d'expérience (REX), que nous avons donc suffisamment de certitudes pour le considérer comme vraisemblable pour l'ensemble de nos produits ou non.

En statistiques, lorsque l'on a besoin de définir un critère d'adéquation entre un modèle de fiabilité prévisionnel par rapport aux informations recueillies en exploitation, on peut passer par l'utilisation d'un test d'adéquation, le test d'hypothèses.

« *Qu'est-ce qu'un test d'hypothèses ?* »

Très souvent, un problème de décision consiste à trancher, au vu d'observations, entre une hypothèse ou une autre.

Par exemple, lorsque l'on tire à « Pile ou Face », qu'on choisit toujours pile, et que l'on a l'impression de perdre un peu trop souvent, ou de gagner un peu trop souvent, on peut se demander si la pièce est truquée ou si elle ne l'est pas.

L'une de ces hypothèse sera appelée hypothèse nulle (notée  $H_0$ ), et l'autre hypothèse qu'on lui oppose sera dite « hypothèse alternative » (notée  $H_1$ ).

En général, une et une seule de ces deux hypothèses peut être vraie. Un test d'hypothèses est donc une procédure qui permet de choisir entre ces deux hypothèses.

Lors de cette procédure, on peut noter deux types d'erreurs de décision possible :

- l'**erreur de première espèce**, qui est de rejeter l'hypothèse  $H_0$  à tort
- l'**erreur de deuxième espèce**, qui est de rejeter l'hypothèse  $H_1$  à tort.

L'idéal serait évidemment de trouver une procédure qui minimise les deux risques d'erreurs en même temps. Malheureusement, on peut démontrer qu'ils varient en sens inverse, c'est-à-dire que toute procédure diminuant l'un va en général, augmenter l'autre et réciproquement.

Dans notre exemple, on se demande si la pièce de monnaie est truquée lors de notre tirage. On peut donc poser les hypothèses suivantes :

-H0 : La pièce de monnaie n'est pas truquée.

-H1 : La pièce de monnaie est truquée.

On pose nos hypothèses ainsi car on veut établir une certaine « présomption d'innocence ». De même que dans un procès d'assises, il y a présomption d'innocence, en théorie des tests, il y a **présomption de H0**, c'est-à-dire qu'il faut montrer que H0 est peu probable pour décider H1.

Par contre accepter H0 peut venir du fait que soit H0 est effectivement vraie, soit on n'a pas les moyens nécessaires de prouver H1. (Ex: pas assez d'observations).

On note  $\alpha$ , la probabilité de l'erreur de première espèce, et  $\beta$  la probabilité de l'erreur de deuxième espèce.

Dans la pratique des tests statistiques, il est de règle de **fixer**  $\alpha$  (ex : 5%, 1%, 10%) car c'est le risque que l'on veut contrôler, celui que l'on est prêt à prendre en rejetant à tort H0, c'est-à-dire le **niveau de risque** que l'on s'accorde quant à l'affirmation des décisions que l'on va prendre par la suite. Ce choix est basé sur la perception que l'on a de la gravité des conséquences d'un rejet injustifié de H0.

Pour mettre en relation la décision prise avec la réalité de la situation (dont on ne peut jamais être sûr à l'avance, voir ne jamais être sûr du tout selon les cas), on pose le tableau suivant :

Décision \ Réalité	H0 : La pièce de monnaie n'est pas truquée.	H1 : La pièce de monnaie est truquée.
H0 : La pièce de monnaie n'est pas truquée.	Probabilité = $1-\alpha$	Probabilité = $\beta$
H1 : La pièce de monnaie est truquée.	Probabilité = $\alpha$	Probabilité = $1-\beta$

Effectivement, on peut maintenant mieux distinguer tous les cas qui peuvent arriver lors d'une prise de décision pareille.

Il devient évident que l'on peut définir les probabilités de ces différents cas de figure comme dans le tableau précédent.

En effet, si on part du principe qu'il y a forcément une réalité absolue, c'est que l'autre cas de figure n'est donc pas possible.

Donc si la réalité est «  $H_0$  : La pièce de monnaie n'est pas truquée. », c'est qu'il ne peut en être autrement, et qu'il ne tient plus qu'à l'homme concerné de prendre la bonne décision ou non. Il en ressort donc que sa probabilité de prendre la mauvaise décision est égale à «  $\alpha$  » (l'erreur de première espèce), et que du coup sa probabilité de prendre la bonne décision est égale à «  $1-\alpha$  ».

De même, si la réalité est «  $H_1$  : La pièce de monnaie est truquée. », il ne peut là non plus en être autrement, et là encore il ne tient plus qu'à l'homme concerné de prendre la bonne décision ou non. Il en ressort donc que sa probabilité de prendre la mauvaise décision est dans ce cas égale à «  $\beta$  » (l'erreur de deuxième espèce), et que du coup, sa probabilité de prendre la bonne décision est égale à «  $1-\beta$  ».

Cette dernière probabilité «  $1-\beta$  » est appelée **la puissance du test**. En effet, vu que nous avons préalablement posé notre test de façon à ce que la probabilité de l'erreur de première espèce «  $\alpha$  » soit la probabilité de déclarer à tort que la pièce est truquée, il va en découler que **la puissance du test** «  $1-\beta$  » soit la probabilité de « **détecter** » à **raison** que la pièce est bien truquée.

*« Comment peut-on maintenant exprimer ce test d'hypothèses de manière statistique ? »*

Si la pièce n'est effectivement pas truquée et que l'on réalise  $N$  tirages ( $N$  assez grand), on devrait normalement obtenir un nombre de réalisations de « Pile » et de « Face » plus ou moins égales, car la probabilité d'obtenir l'une ou l'autre des situations est de 50%.



Par exemple,  $N = 100$ , on lance 100 fois la pièce et on note  $X$  la variable qui compte le nombre de fois que l'on obtient « pile ».

Si on réalise un test **bilatéral**, on peut donc adopter la règle de décision suivante :

$H_0$  : la pièce n'est pas truquée est,

- acceptée si  $X \in [40,60]$ ,
- rejetée si  $X \notin [40,60]$ .

Dans le cadre de ce test, on a :

$$P(\text{rejeter } H_0 / H_0 \text{ vrai}) = P(X \notin [40, 60]) = \alpha \approx 0,05$$

et donc

$$P(\text{accepter } H_0 / H_0 \text{ vrai}) = P(X \in [40, 60]) = 1 - \alpha \approx 0,95.$$

Il s'agit donc d'un test où l'on s'est fixé  $\alpha = 0,05$ .

→ Si la valeur de  $X$  que l'on obtient se trouve dans la zone de rejet (région critique, c'est-à-dire en dehors de  $[40, 60]$ ), on dira que la statistique  $X$  observée est statistiquement significative au seuil  $\alpha$ . Cette valeur est anormalement élevée ou basse et ne permet pas d'accepter  $H_0$ . On rejette  $H_0$ . On ne peut pas accepter que la pièce de monnaie ne soit pas truquée.

→ Si la valeur de  $X$  que l'on obtient se trouve dans la zone d'acceptation (c'est-à-dire en à l'intérieur de  $[40, 60]$ ), on dira que la statistique  $X$  observée n'est pas significative au seuil  $\alpha$ . Cette valeur est imputable aux fluctuations d'échantillonnage. On accepte  $H_0$ . On ne peut pas rejeter le fait que la pièce de monnaie ne soit pas truquée.

**Cas du REX et des modèles de fiabilité prévisionnelle :**

Maintenant que la définition d'un test d'hypothèse est plus clairement définie, nous pouvons revenir à son utilisation dans le cadre de l'adéquation entre notre REX et un modèle de fiabilité prévisionnelle.

On pose les hypothèses suivantes :

- H0 : Le modèle de fiabilité prévisionnelle est compatible du REX.
- H1 : Le modèle de fiabilité prévisionnelle n'est pas compatible du REX.

Ce qui donne le tableau suivant :

Décision \ Réalité	H0 : Le modèle de fiabilité prévisionnelle est compatible du REX.	H1 : Le modèle de fiabilité prévisionnelle n'est pas compatible du REX.
H0 : Le modèle de fiabilité prévisionnelle est compatible du REX.	Probabilité = $1-\alpha$	Probabilité = $\beta$ (erreur deuxième espèce)
H1 : Le modèle de fiabilité prévisionnelle n'est pas compatible du REX.	Probabilité = $\alpha$ (erreur première espèce)	Probabilité = $1-\beta$ (puissance test)

**La puissance** de ce test nous permet donc de détecter les cas où l'on rejette la compatibilité entre le modèle de fiabilité prévisionnelle et le REX à raison. C'est-à-dire qu'elle nous indique **la probabilité de détecter** les cas où il y a « réellement » incompatibilité (la différence est significative et détectable).

Dans le cadre de notre test, **l'estimateur du taux de défaillance** pour les informations recueillies (REX) est donné par :

$$\hat{\lambda} = \frac{\text{Nombre de pannes}}{\sum \text{Hrs}} = \frac{r}{T}$$

Donc « r » est le nombre de pannes (défaillances) que l'on a observées sur nos données durant notre expérience et « T » est la somme des heures cumulées.

Il s'agit des deux paramètres d'entrées que l'on renseigne dans la zone suivante :

*Retour d'expérience (REX)*

Nombre de pannes observées

Somme des heures

Nous sommes dans un cadre où nous définirons nos données comme censurées à droite (Type I), avec la possibilité de remplacement c'est-à-dire que dès qu'un produit observé est en panne, il est soit immédiatement remplacé soit immédiatement réparé.

Il s'agit donc d'essais stoppés après un temps «  $t_0$  » fixé à l'avance, quel que soit le **nombre « r » de pannes observées.**

Si on a « N » **entités en observation** et que l'on ne considèrerait pas la possibilité de remplacement en cas de panne, la somme des heures cumulées « T », s'exprimerait sous la forme :

$$T = t_1 + t_2 + \dots + t_r + (N - r)t_0$$

Cependant, nous sommes dans le cadre de remplacement ou réparation immédiate en cas de panne (le temps de remplacement ou réparation est négligeable dans le calcul), donc **la somme des heures cumulées « T »**, s'exprime simplement sous la forme :

$$T = N * t_0$$

On peut donc calculer (grâce à un certain nombre de règles statistiques importantes), un intervalle de confiance en fonction de  $\alpha$  pour notre taux de défaillance observé (avec un niveau de risque  $\alpha$  qu'on s'est fixé et donc un  $NC=1-\alpha$  niveau de confiance que l'on s'est fixé) :

$$[\hat{\lambda}_{min}; \hat{\lambda}_{max}] \text{ tel que } P(\text{accepter } H_0 | H_0 \text{ vrai}) = P(\lambda_{FIDES} \in [\hat{\lambda}_{min}; \hat{\lambda}_{max}]) = NC = 1 - \alpha$$

Il s'agit du niveau de risque  $\alpha$  que vous avez renseigné comme paramètre d'entrée :

*Précision du test*

Niveau de risque :

10%	▼
10%	
5%	
1%	
0.5%	
0.1%	

On dit qu'un retour d'expérience (REX) est **significatif** si le taux de défaillance prévisionnel donné par la méthodologie FIDES **n'est pas compris** dans l'intervalle de confiance du taux de défaillance observé.

J'affirme donc que, si le taux de défaillance prévisionnel «  $\lambda_{\text{FIDES}}$  » est en dehors de l'intervalle de confiance du REX, le REX est significatif en ce sens que l'on peut rejeter l'hypothèse comme quoi le taux de défaillance prévisionnel est compatible du REX.

Votre «  $\lambda_{\text{FIDES}}$  » se trouve en dehors de l'intervalle obtenu grâce à votre REX, c'est la raison pour laquelle vous obtenez le résultat suivant :

**Résultat :**

Votre taux de défaillance prévisionnel donné par Fides n'est pas compatible avec votre retour d'expérience

[Plus de détails](#)

Votre modèle de fiabilité prévisionnelle FIDES ne répond pas au critère permettant la compatibilité avec le REX. Nous ne pouvons pas aller plus loin dans l'expression de la qualité de la compatibilité entre votre REX et votre taux de défaillance prévisionnel donné par FIDES.