

# TESTS D'HYPOTHÈSES

---

ED – PASS Lyon Est – 2024-2025

Dr. Nicolas ROMAIN-SCELLE

# Question 1 (QCM) (*Biostatistique, 2009*)

Concernant le résultat des tests d'hypothèses, quels sont les énoncés vrais ?

- A. Si le test est statistiquement non significatif, on ne rejette pas  $H_0$
- B. Si le test est statistiquement significatif, on ne rejette pas  $H_0$
- C. Si le test est statistiquement significatif, on est certain que  $H_0$  est fausse
- D. Si le test est statistiquement significatif, on est certain que  $H_1$  est vraie
- E. Si le test est statistiquement non significatif, on est certain que  $H_1$  est fausse

# Question 1 (QCM)

Concernant le résultat des tests d'hypothèses, quels sont les énoncés vrais ?

- A. Si le test est statistiquement non significatif, on ne rejette pas  $H_0$
- B. Si le test est statistiquement significatif, on ne rejette pas  $H_0$
- C. Si le test est statistiquement significatif, on est certain que  $H_0$  est fausse
- D. Si le test est statistiquement significatif, on est certain que  $H_1$  est vraie
- E. Si le test est statistiquement non significatif, on est certain que  $H_1$  est fausse

# Question 1 (QCM)

**La procédure de test ne vise qu'un seul objectif : prouver, avec un risque  $\alpha$  prédéterminé de faire erreur, que l'hypothèse nulle est fausse.**

Aussi :

- On n'acquiert jamais de certitude sur la véracité de  $H_0$  et  $H_1$
- On ne prouve jamais  $H_0$
- On ne prouve pas non plus  $H_1$  : au mieux, on admet l'existence d'une différence, sans prouver que la différence estimée dans l'échantillon est la vraie différence théorique

## Question 2 (QCM)

Concernant les statistiques de test, quels sont les énoncés vrais ?

- A. Une statistique de test est une variable aléatoire
- B. Une statistique de test nécessite une hypothèse nulle pour être calculable
- C. Une statistique de test suit une loi de probabilité
- D. Une statistique de test estime une probabilité
- E. Une statistique de test est une mesure

## Question 2 (QCM)

Concernant les statistiques de test, quels sont les énoncés vrais ?

- A. Une statistique de test est une variable aléatoire
- B. Une statistique de test nécessite une hypothèse nulle pour être calculable
- C. Une statistique de test suit une loi de probabilité
- D. Une statistique de test estime une probabilité
- E. Une statistique de test est une mesure

# La statistique de test

- De façon générale, on note  $S \sim \mathcal{L}(\theta)$  une statistique de test, suivant une loi  $\mathcal{L}$  non spécifiée de paramètre  $\theta$ .
- $\theta$  correspond à une vraie valeur, **non connue**, à laquelle on va comparer l'estimation issue de l'expérience  $t$
- Sous  $H_0: \theta = \theta_0$  avec  $\theta_0$  une valeur **connue ou estimable**
- Donc  $s_{H_0}^{obs} = f(\theta_0, t)$  est désormais calculable :  $\theta_0$  est connu et  $T$  permet d'obtenir l'estimation  $t$  depuis l'échantillon
- $S$  est **par définition** une mesure : elle est construite pour mesurer l'écart entre le résultat d'une réalisation de l'expérience sous  $H_0$  et l'espérance de l'expérience sous  $H_0$ . Elle n'est pas une probabilité.

## Question 3 (QCM)

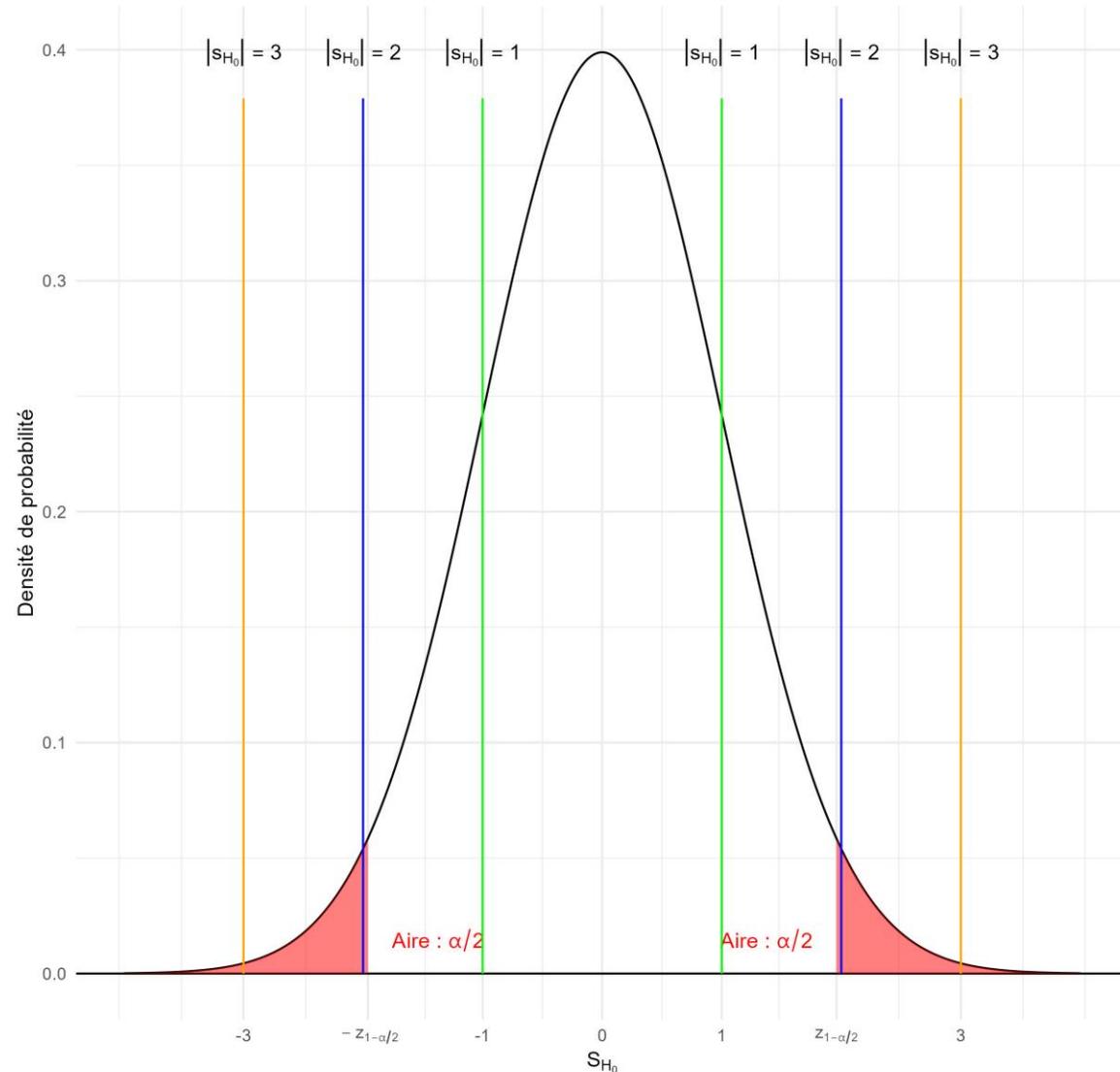
- Vous participez à un jeu de pile ou face contre un adversaire. L'adversaire vous propose trois pièces (verte, bleue, orange) pour jouer. On définit  $X$  la variable aléatoire suivant une loi de Bernoulli de paramètre  $\pi$  caractérisant le résultat d'un lancer :  $X(Pile) = 1$  et  $X(Face) = 0$ .
- Pour chaque pièce, vous réalisez l'expérience suivante: la pièce est lancée 10 fois et vous réalisez un test statistique en prenant les hypothèses suivantes :  $H_0: \pi = 0,5$  et  $H_1: \pi \neq 0,5$
- Vous obtenez les résultats suivants pour les valeurs de la statistique de test:  $s_v = 1$  pour la pièce verte,  $s_b = 2$  pour la pièce bleue et  $s_o = 3$  pour la pièce orange. La figure suivante représente la distribution de la statistique de test  $S$  sur laquelle vous avez ajouté les 3 valeurs  $s_b$ ,  $s_v$  et  $s_o$  ainsi que le risque alpha :

## Question 3 (QCM)

Quels énoncés sont corrects parmi les suivants :

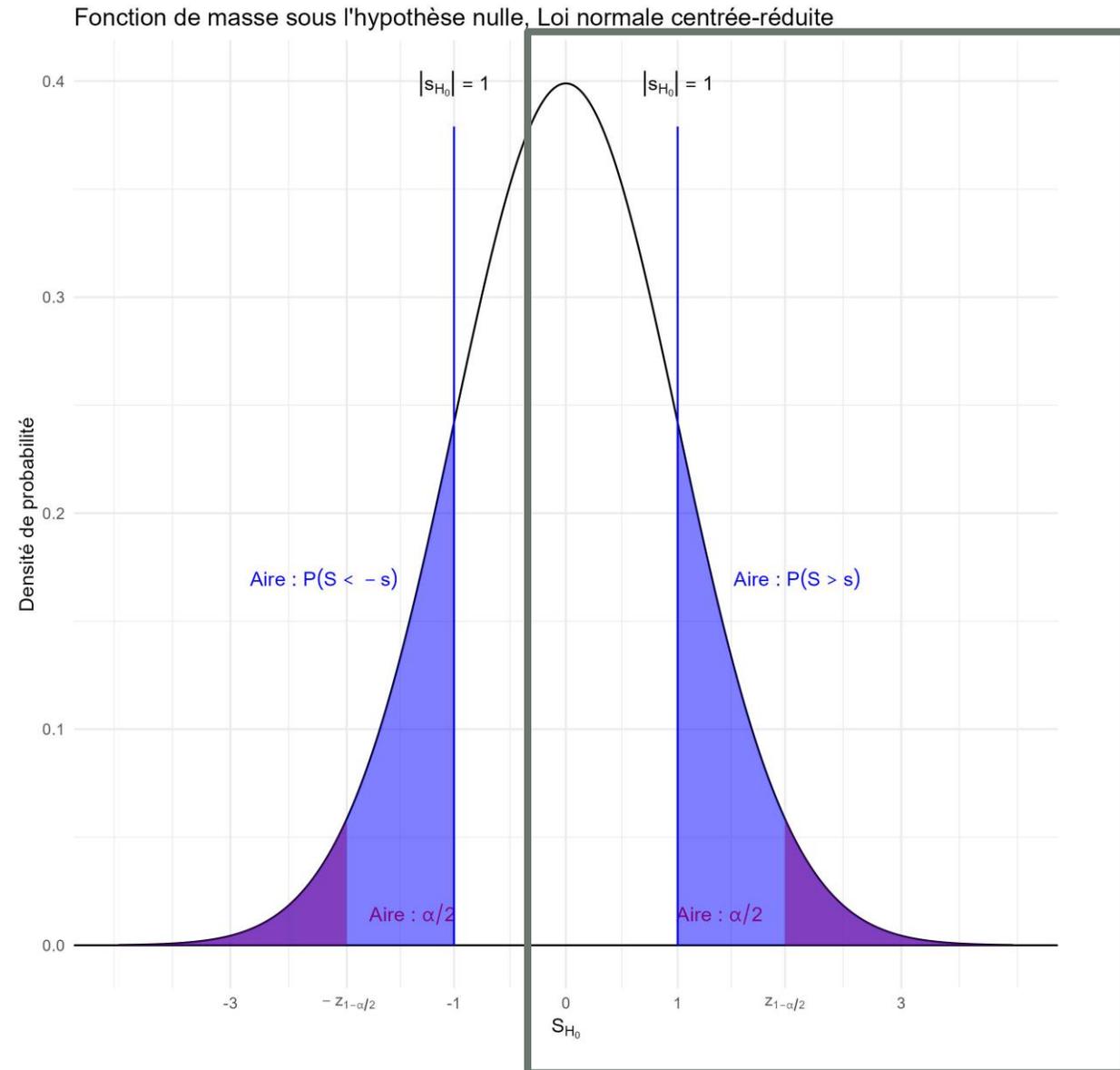
- A. La p-value du test sur la pièce verte est supérieure au seuil de rejet
- B. La p-value du test sur la pièce bleue est supérieure au risque  $\alpha$
- C. La p-value du test sur la pièce orange est inférieure à celle du test sur la pièce verte
- D. La pièce orange est la plus désirable pour obtenir un jeu équilibré
- E. La pièce bleue est la plus désirable pour obtenir un jeu équilibré

Fonction de masse sous l'hypothèse nulle, Loi normale centrée-réduite



# Question 3 (QCM)

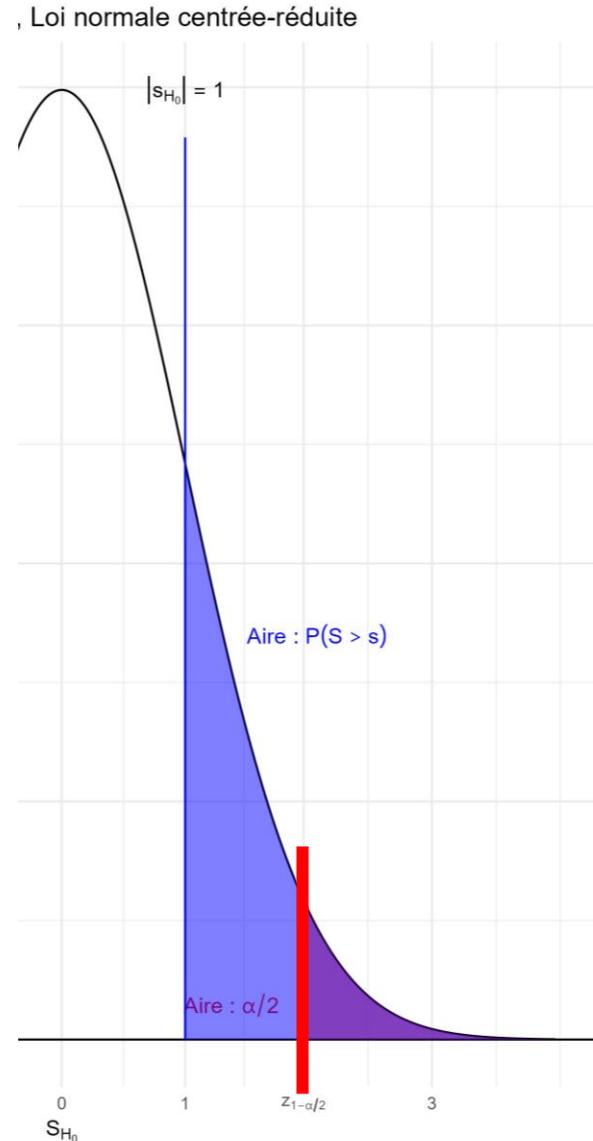
Rappel (exemple à droite)



## Question 3 (QCM)

Rappel (exemple à droite)

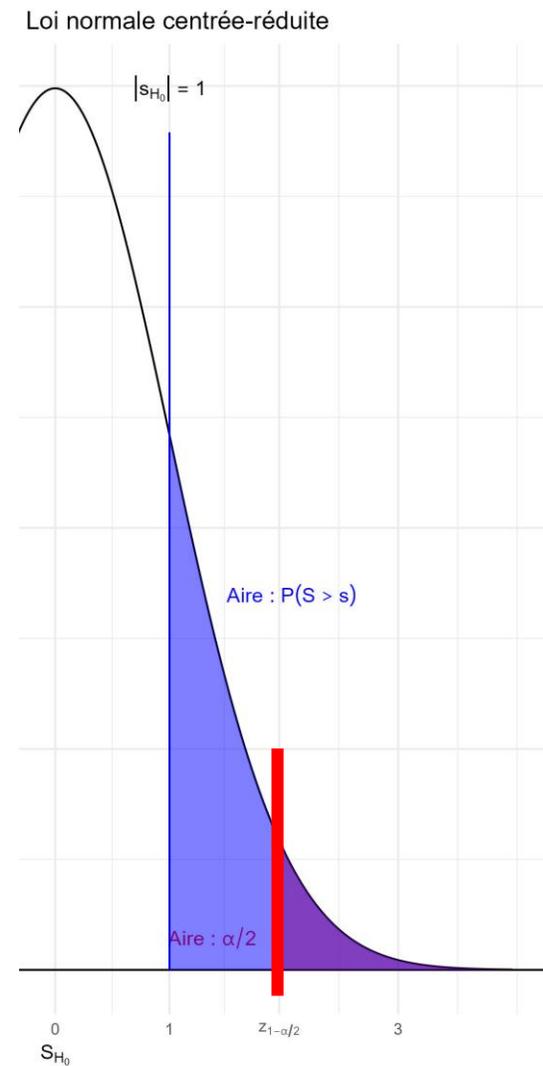
- Le risque  $\alpha$  est **l'aire** définie par l'axe des abscisses en bas, la densité de probabilité en haut, et le **seuil de rejet** à gauche
- La p-value est **l'aire** définie comme le risque  $\alpha$  en haut et en bas, mais par la **statistique de test calculée** à gauche



## Question 3 (QCM)

Rappel (exemple à droite)

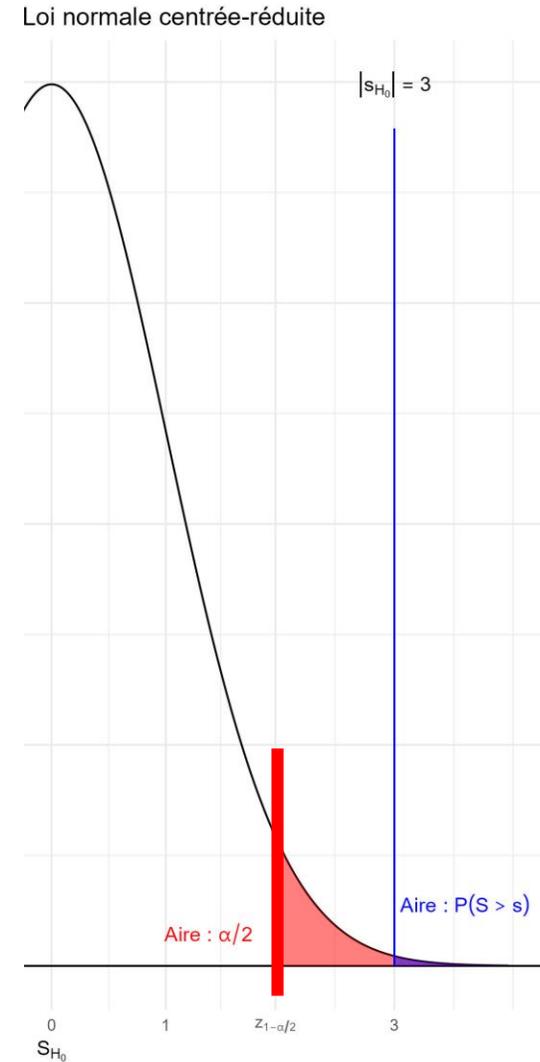
- Si la statistique de test est **inférieure en valeur absolue** au seuil de rejet, alors **la p-value est supérieure au risque  $\alpha$**



## Question 3 (QCM)

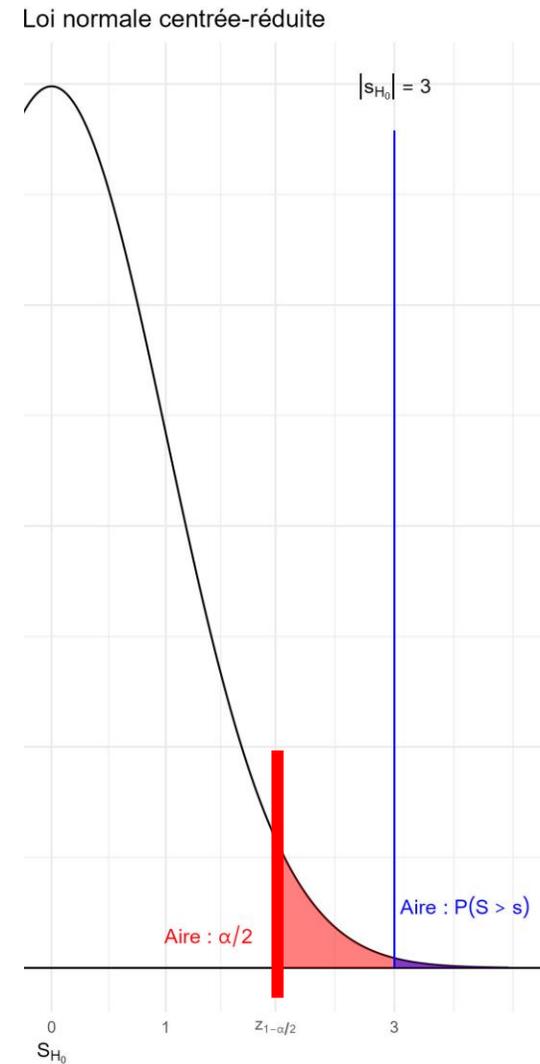
Rappel (exemple à droite)

- Si la statistique de test est **inférieure en valeur absolue** au seuil de rejet, alors **la p-value est supérieure au risque  $\alpha$**
- Si la statistique de test est **supérieure en valeur absolue** au seuil de rejet, alors **la p-value est inférieure au risque  $\alpha$**



## Question 3 (QCM)

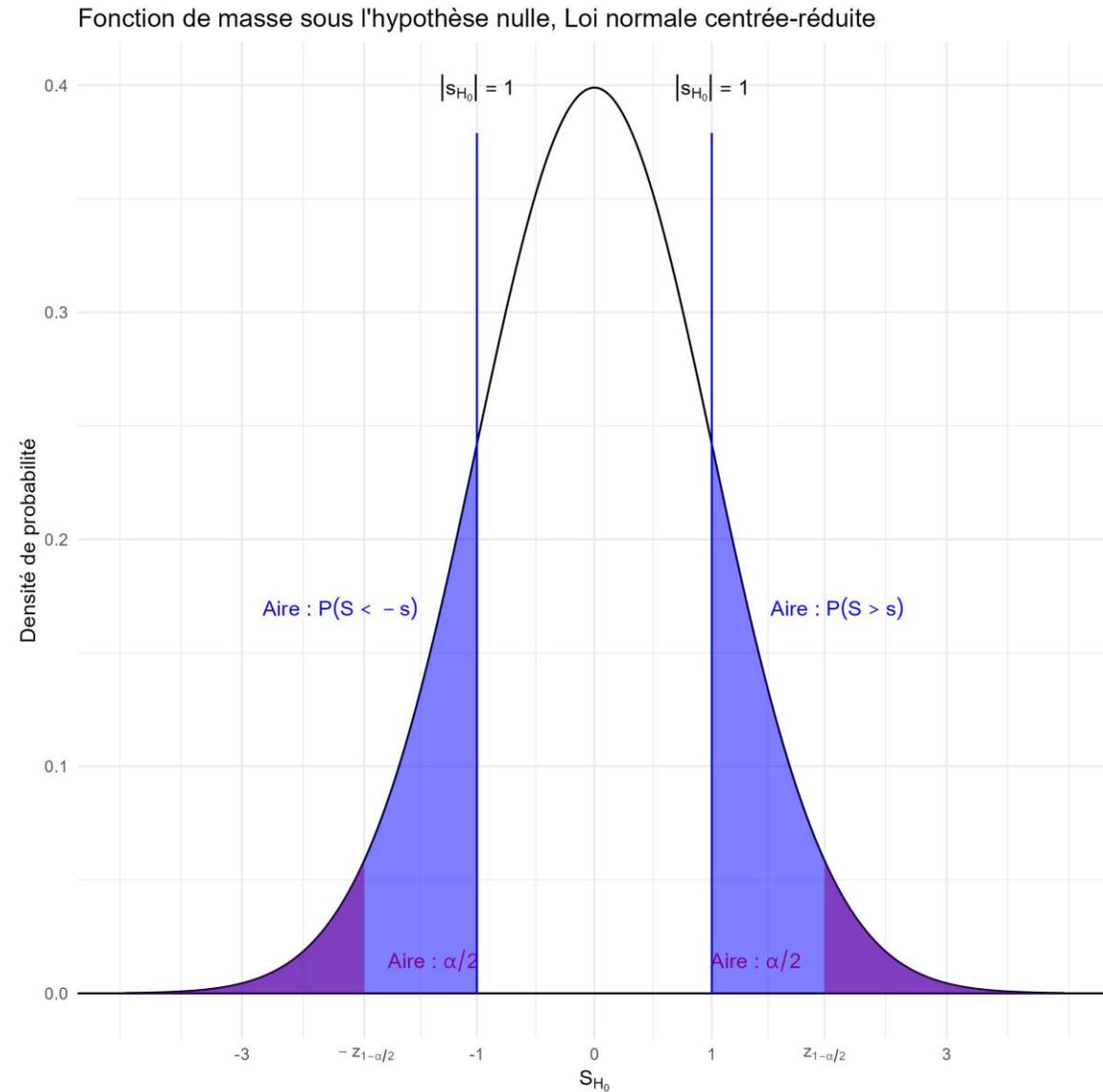
On ne compare pas p-value avec le seuil de rejet, ou la statistique de test avec le risque  $\alpha$  directement



## Question 3 (QCM)

$$|s_{H_0}| = 1$$

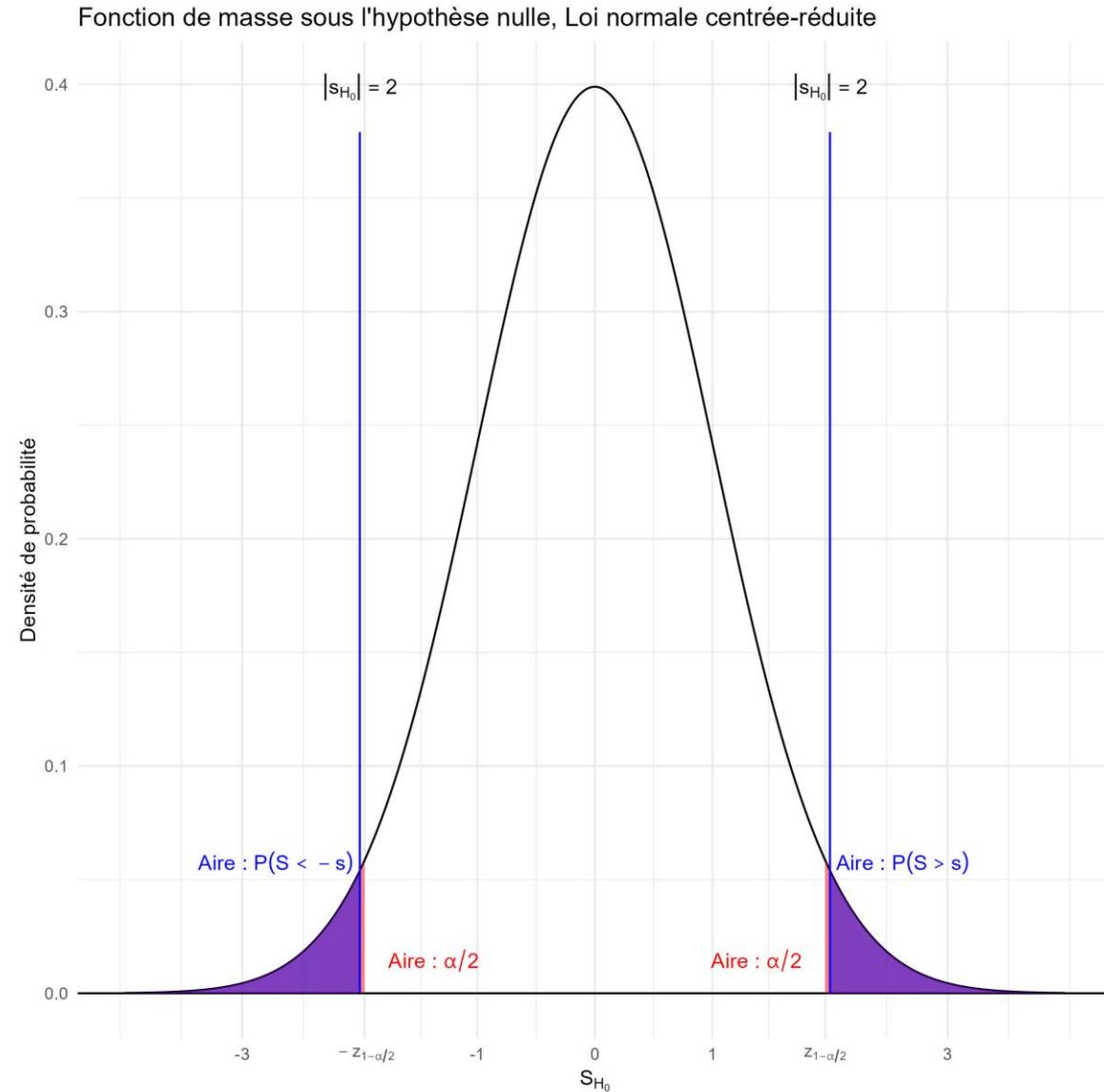
- La p-value (aire bleue) est supérieure au risque  $\alpha$  (aire rouge)
- La statistique de test est inférieure en valeur absolue au seuil de rejet
- **Le test n'est pas concluant** : on ne rejette pas l'hypothèse nulle



## Question 3 (QCM)

$$|s_{H_0}| = 2$$

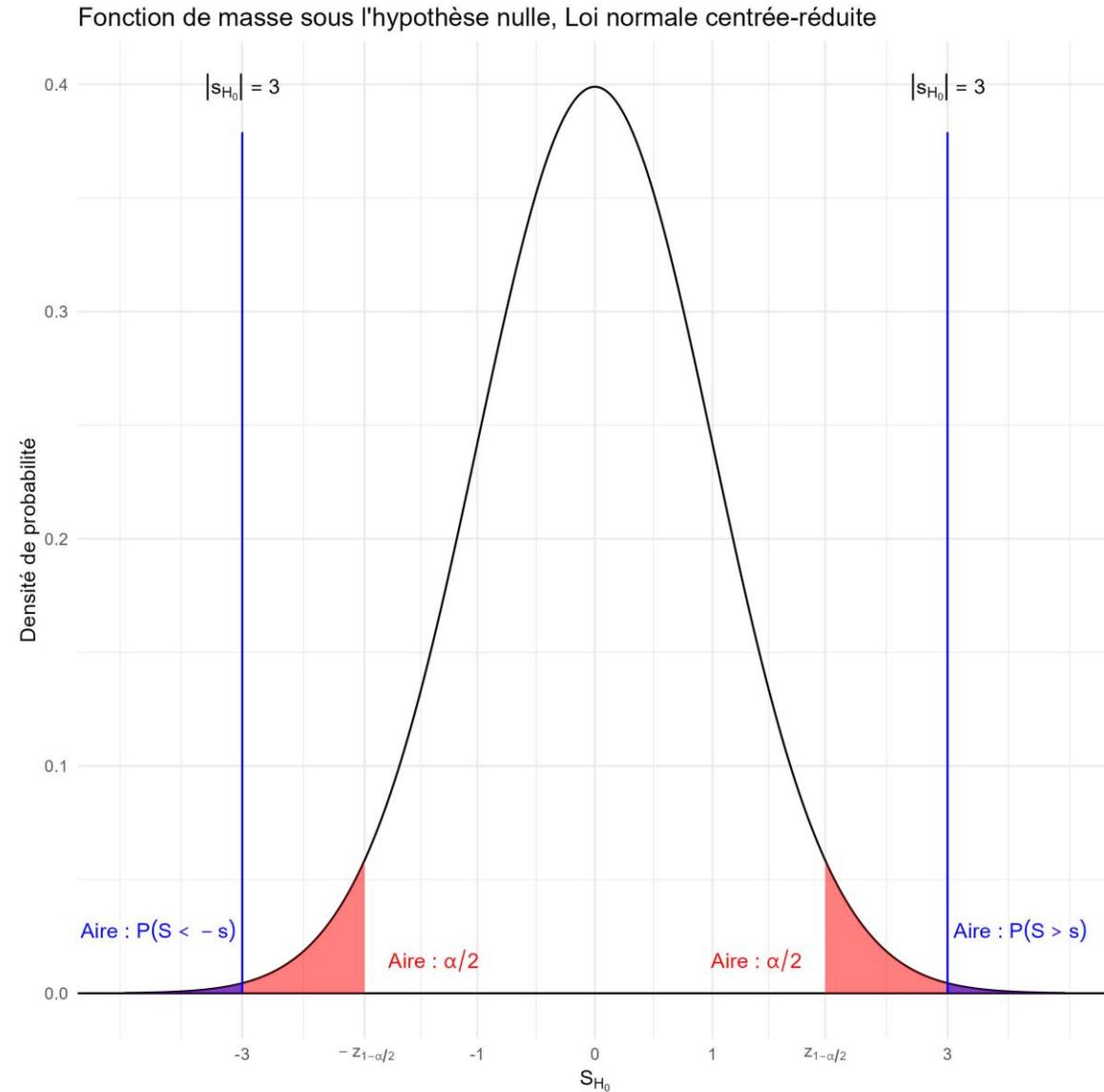
- La p-value (aire bleue) est inférieure au risque  $\alpha$  (aire rouge)
- La statistique de test est supérieure en valeur absolue au seuil de rejet
- **Le test est concluant :** on rejette l'hypothèse nulle avec un risque  $\alpha$  de faire erreur



## Question 3 (QCM)

$$|s_{H_0}| = 3$$

- La p-value (aire bleue) est inférieure au risque  $\alpha$  (aire rouge)
- La statistique de test est supérieure en valeur absolue au seuil de rejet
- **Le test est concluant :** on rejette l'hypothèse nulle avec un risque  $\alpha$  de faire erreur

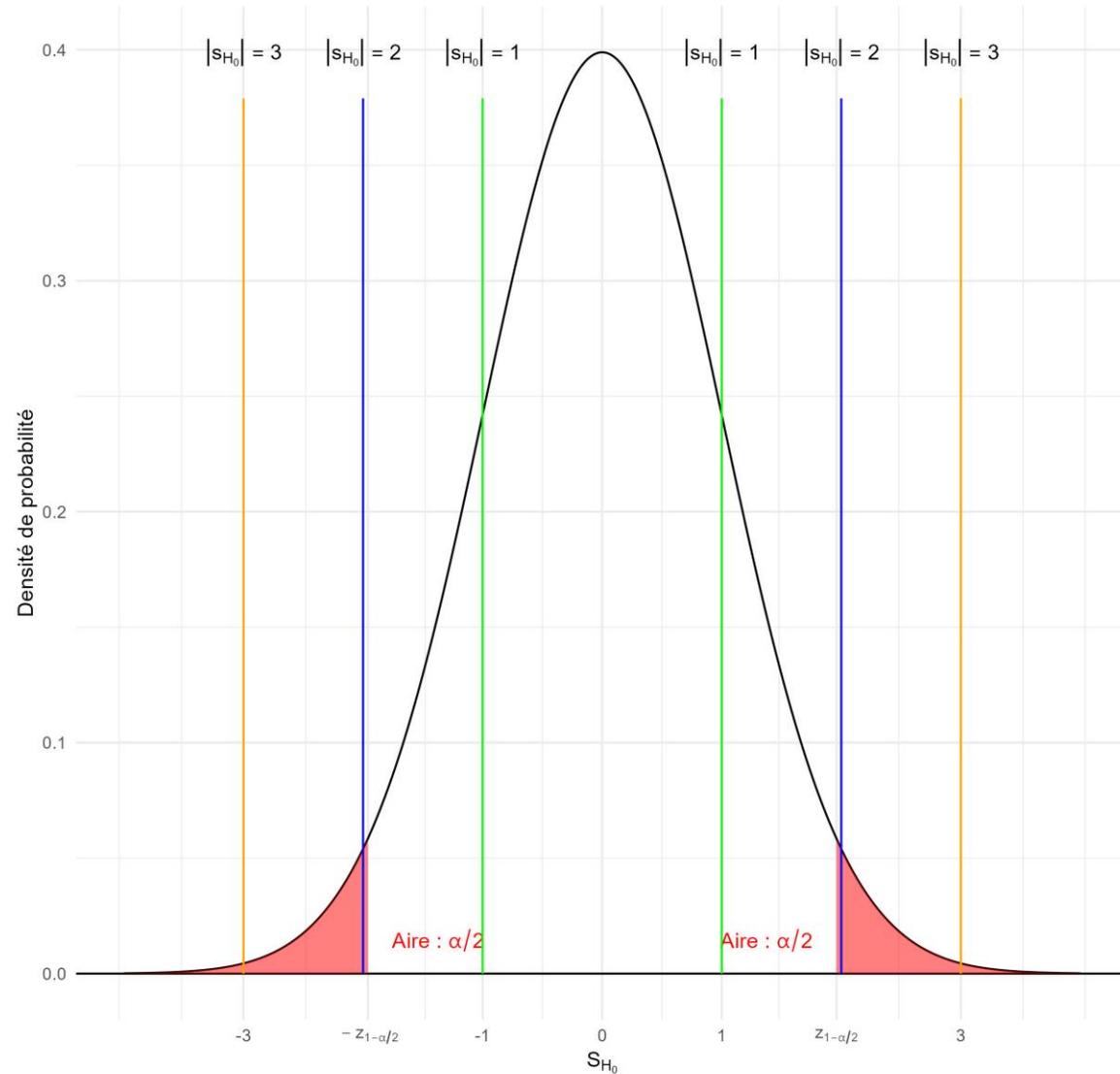


## Question 3 (QCM)

Quels énoncés sont corrects parmi les suivants :

- A. La p-value du test sur la pièce verte est supérieur au seuil de rejet
- B. La p-value du test sur la pièce bleue est supérieure au risque  $\alpha$
- C. La p-value du test sur la pièce orange est inférieure à celle du test sur la pièce verte
- D. La pièce orange est la plus désirable pour obtenir un jeu équilibré
- E. La pièce bleue est la plus désirable pour obtenir un jeu équilibré

Fonction de masse sous l'hypothèse nulle, Loi normale centrée-réduite



## Question 4 (QRU)

- Vous êtes sollicités pour réaliser un test du  $\chi^2$  sur la table de contingence suivante, croisant deux variables catégorielles X et Y mesurées sur un échantillon. Donnez le nombre de degrés de liberté du test.

Observé	$y = A$	$y = B$	$y = C$	
$x = 1$	39	9	39	87
$x = 2$	37	26	29	92
$x = 3$	19	4	45	68
$x = 4$	17	49	44	110
$x = 5$	18	8	49	75
	130	96	206	432

## Question 4 (QRU)

Les **distributions marginales** de  $x$  et  $y$  sont fixes (cellules rouges)

La **distribution conditionnelle** de  $x$  sachant  $y$  (et de  $y$  sachant  $x$ ) est aléatoire.

On cherche à savoir à partir de combien de cellules bleues remplies l'ensemble du tableau est connu

Observé	$y = A$	$y = B$	$y = C$	
$x = 1$	?	?	?	87
$x = 2$	?	?	?	92
$x = 3$	?	?	?	68
$x = 4$	?	?	?	110
$x = 5$	?	?	?	75
	130	96	206	432

## Question 4 (QRU)

- Le test du  $\chi^2$  pour cette table vise à rechercher une différence de distribution entre 2 variables catégorielles à plus de 2 modalités.
- Sous  $H_0$ , on va varier les effectifs **conditionnels** uniquement, les effectifs **marginaux** étant considérés comme fixes.

Observé	$y = A$	$y = B$	$y = C$	
$x = 1$	39	9	39	87
$x = 2$	37	26	29	92
$x = 3$	19	4	45	68
$x = 4$	17	49	44	110
$x = 5$	18	8	49	75
	130	96	206	432

# Question 4 (QRU)

- On doit déterminer le nombre de modalités de chaque variable
  - Pour X :  $I = 5$  modalités {1,2,3,4,5} => **Nombre de lignes** (hors total)
  - Pour Y :  $J = 3$  modalités {A,B,C} => **Nombre de colonnes** (hors total)
- On détermine également le nombre de **contraintes** s'exerçant sur chaque variable :
  - Sur X : 1 contrainte : **total des effectifs par modalités** (dernière colonne)
  - Sur Y : 1 contrainte : **total des effectifs par modalités** (dernière ligne)

## Question 4 (QRU)

- Exemple : pour la première ligne ( $x = 1$ )
  - L'égalité  $O_{x=1,y=A} + O_{x=1,y=B} + O_{x=1,y=C} = O_{x=1} = 39$  est toujours vraie
  - Donc :  $O_{x=1,y=A} = 39 - O_{x=1,y=B} + O_{x=1,y=C}$
- L'extension de cette logique pour une table de contingence du test du  $\chi^2$  conduit à la formule :
$$(I - 1) \times (J - 1)$$
- Avec  $I = 5$  et  $J = 3$  :  $(5 - 1) \times (3 - 1) = \mathbf{8 \text{ ddl}}$

## Question 5 (QCM)

A l'issue d'un essai clinique visant à comparer l'effet d'un nouveau traitement A à un traitement de référence B sur la mortalité à 1 an, vous disposez des résultats suivants : 53 patients sous A sont décédés, 68 sous B sont décédés, avec 144 patients traités par A et 147 patients traités par B.

On fixe  $\alpha = 0,025$ ,  $H_0: \pi_A = \pi_B = \pi_0$  et  $H_1: \pi_A \neq \pi_B \neq \pi_0$

Quels sont les énoncés corrects parmi les suivants ?

- A. L'écart entre la statistique du test du Chi-deux calculée et le seuil de rejet associé est supérieur à 1 en valeur absolue
- B. L'écart entre la statistique du test du Chi-deux calculée et le seuil de rejet associé est supérieur à 3 en valeur absolue
- C. Sous l'hypothèse nulle, on attend entre 59 et 60 décès dans le groupe A
- D. On rejette l'hypothèse nulle au risque  $\alpha$
- E. On ne rejette pas l'hypothèse nulle au risque  $\alpha$

# Question 5 (QCM)

On pose les hypothèses :

$$H_0: \pi_A = \pi_B = \pi_0$$

$$H_1: \pi_A \neq \pi_B \neq \pi_0$$

Avec  $\pi_i$  la vraie proportion de décédés dans les groupes.

# Sous l'hypothèse nulle

Obs.	A	B	
Décédé	53	68	121
Vivant	91	79	170
	144	147	291

Sous $H_0$	A	B	
Décédé	<b>59,9</b>	61,1	121
Vivant	84,1	85,9	170
	144	147	291

# Conditions d'applications

Sous $H_0$	A	B	
Décédé	$n_A f_0 = 59,9 > 5$	$n_B f_0 = 61,1 > 5$	121
Vivant	$n_A(1 - f_0) = 84,1 > 5$	$n_B(1 - f_0) = 85,9 > 5$	170
	144	147	$n = 291$

Les conditions d'application du test du **Chi-deux** sont vérifiées dans la table de contingence **sous  $H_0$**

# Statistique de test

$$S_{H_0}^{obs} = \sum_{i=1}^{I=2} \sum_{j=1}^{J=2} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \cong 2,30$$

- $\alpha = 0,025$  donc  $q_{1-\alpha}^{\chi^2} \cong 5,02$
- **On ne rejette pas  $H_0$**
- $5,02 - 2,30 > 1$  donc **A vraie** mais  $5,02 - 2,30 < 3$  donc **B fausse**

## Question 5 (QCM)

A l'issue d'un essai clinique visant à comparer l'effet un nouveau traitement A à un traitement de référence B sur la mortalité à 1 an, vous disposez des résultats suivants : 53 patients sous A sont décédés, 68 sous B sont décédés, avec 144 patients traités par A et 147 patients traités par B.

On fixe  $\alpha = 0,025$ ,  $H_0: \pi_A = \pi_B = \pi_0$  et  $H_1: \pi_A \neq \pi_B \neq \pi_0$

Quels sont les énoncés corrects parmi les suivants ?

- A. L'écart entre la statistique du test du Chi-deux et le seuil de rejet associé est supérieur à 1 en valeur absolue
- B. L'écart entre la statistique du test du Chi-deux et le seuil de rejet associé est supérieur à 3 en valeur absolue
- C. Sous l'hypothèse nulle, on attend entre 59 et 60 décès dans le groupe A
- D. On rejette l'hypothèse nulle au risque  $\alpha$
- E. On ne rejette pas l'hypothèse nulle au risque  $\alpha$