Etude comparative de trois fonctions de transformation normalisante d'une proportion (Anscombe, 1948 – Freeman-Tukey, 1950 – Chanter 1975)

Louis Laurencelle ^a

^aUniversité du Québec à Trois-Rivières

Abstract ■ This essay on proportions explores the statistical properties of three normalizing functions of a proportion based on Fisher's arcsine transformation (Anscombe, 1948; Freeman & Tukey, 1950; Chanter, 1975), namely their bias, variance and precision, quasi-normality. The use of these functions in statistical testing procedures is exemplified for the case of one proportion versus a given parametric value and that of the difference between two proportions, with a check on the procedures' accuracy and power. Conventional test procedures as well as exact calculation solutions are provided for comparison. An abundant graphic file accompanies the article, thanks to which the reader can make his own idea on the arguments and demonstrations here proposed.

Keywords ■ Proportions, Transformation arcsinus, Propriétés statistiques, Test sur une proportion, Test de différence entre deux proportions ■ Proportions, Arcsine transformation, Statistical properties, Test on one or two proportions..

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

louis.laurencelle@gmail.com

[10.20982/tqmp.17.3.p166](https://doi.org/10.20982/tqmp.17.3.p166)

Introduction

Depuis ANSCOMBE (1948) et sa fonction de transformation, d'autres (p. ex. FREEMAN & TUKEY, 1950; CHANTER, 1975; RÜCKER, SCHWARZER, CARPENTER & OLKIN, 2009; WANG & FINGERT, 2012) ont étudié le problème de la transformation, normalisante ou non, d'une proportion, ce en proposant une, voire plusieurs fonctions. On a comparé les fonctions proposées par chacun, soit théoriquement, soit par simulation Monte Carlo et sous des conditions di-

verses, notamment celles concernant des proportions très basses. Certains ont conclu en faveur de l'une ou l'autre fonction, le but de tous étant d'optimiser l'analyse des proportions, notamment trouver une fonction de transformation $f(x, n)$ qui soit à biais contrôlable, variance stable et distribution normale.

Nous considérons ici la fonction angulaire générique $\sin^{-1}(\sqrt{x/n})$, attribuée à R. A. Fisher, à variance paramétrique constante selon π . Les trois fonctions retenues ici, des variantes de celle de Fisher, sont :

$$\text{Anscombe :} \quad y_{ANS} = \sin^{-1} \sqrt{\frac{x + 3/8}{n + 3/4}}; \quad \text{var}(y) \approx 1/(4n + 2)$$

$$\text{Freeman et Tukey :} \quad y_{F-T/2} = \frac{1}{2} \left(\sin^{-1} \sqrt{\frac{x}{n+1}} + \sin^{-1} \sqrt{\frac{x+1}{n+1}} \right); \quad \text{var}(x) \approx 1/(4n + 2)$$

$$\text{Chanter :} \quad y_{CHA} = \sin^{-1} \sqrt{\frac{1}{4n} + \left(1 - \frac{1}{2n}\right) \frac{x}{n}}; \quad \text{var}(x) \approx 1/4n,$$

la transformation angulaire $\sin^{-1}$ (ou $\arcsin$) étant effectuée en radian.

Noter que la formule présentée par FREEMAN et TUKEY (1950) ne comportait pas la réduction par 2 appliquée souvent par nous (et alors notée $F-T/2$ dans l'article), laquelle permettra ici la comparaison des trois fonctions sur une même échelle, sans interférer aucunement avec ses propriétés statistiques.

Nous examinerons, dans l'ordre, le *biais*, la *variance*, la *précision* (par l'EQM, ou erreur quadratique moyenne), la *normalité*, l'*exactitude* (ou respect du seuil α) et la puissance relatifs aux trois fonctions, pour conclure par un bilan sur leurs mérites comparés. Pour la partie de l'étude concernant l'exactitude et la puissance, via le test de différence entre deux proportions indépendantes, nous mettrons aussi en lice le test z classique assigné à cette tâche de même que notre test à calcul binomial exact (LAURENCELLE, 2005, 2017, 2021b), un calcul plus satisfaisant mais globalement équivalent au test plus simple de LIDDELL (1978).

Enfin, nos estimations et calculs ont été réalisés par programmation sur langage Pascal-Delphi et ont mis en jeu des simulations Monte Carlo, la variable binomiale $x \sim \text{Bin}(\pi, n)$ étant produite soit par élément x_i ($x_i = 1$ si $U \leq \pi$, où $U \sim \text{Unif}(0, 1)$ et $x = \sum_{i=1}^n x_i$), soit directement pour x par la technique du tableau développé (DEVROYE, 1986; LAURENCELLE, 2001) : selon les cas, le nombre d'échantillons aléatoires a été de 50 000 à 1 000 000. D'autres calculs, tels les moments de la variable, ont été effectués exactement selon les probabilités de la distribution binomiale.

Le biais des fonctions

Notation importante. La lettre grecque « π », classiquement associée à la valeur de la demi-circonférence d'un cercle de rayon 1, est aussi employée pour identifier le paramètre de probabilité du processus de Bernoulli, appliqué notamment dans la loi binomiale. Le lecteur est donc prévenu que la notation π servira ici à ce second usage, indiquant la probabilité qu'un événement E_i à deux résultats possibles se produise en faveur du résultat convenu « positif », soit $x_i = 1$ plutôt que $x_i = 0$. La loi binomiale, mettant en jeu n événement, donnera donc lieu à la variable $x = \sum_{i=1}^n x_i$, sa distribution notée $x \sim B(n, \pi)$ et la proportion brute $p = x/n$.

Le biais d'une fonction d'estimation se définit comme l'écart entre l'espérance (ϵ) de cette fonction sous des conditions données et sa valeur théorique ou asymptotique. Soit par exemple $\sum_{j=1}^t (x_j/n)/t = \sum_{j=1}^t p_j/t \rightarrow \epsilon(p_j) = \pi$ lorsque $t \rightarrow \infty$, alors, pour nos fonctions

étudiées, soit $y(x, n)$, le biais asymptotique est simplement la différence $\epsilon[y(x, n)] - \sin^{-1}\sqrt{\pi}$. Au contraire de la proportion simple p pour laquelle $\epsilon(p) - \pi = 0$ pour tout n , ce n'est pas le cas des fonctions angulaires $y(x, n)$, dont les biais *tendent* asymptotiquement vers 0 sur $n \rightarrow \infty$. Il s'agit d'un biais structurel, lié à la fonction de transformation angulaire « y » et qui, dans notre cas, entraîne sous de faibles valeurs de n une surestimation de « y » pour de petites valeurs de π ($< 1/2$) et une sous-estimation correspondante pour de grandes valeurs de π ($> 1/2$). Noter que le problème du biais en espérance n'est ni original, ni spécifique de la fonction angulaire : l'écart-type (s), la corrélation linéaire (r), les centiles et bien d'autres statistiques en sont aussi affectés (KENDALL & STUART, 1977). Ce biais des fonctions « y » selon la taille n , le lecteur le comprendra, n'importe en pratique que lorsque des tailles n différentes sont en jeu : si les tailles n sont égales, alors les différentes valeurs de la fonction sont équitablement comparables, peuvent être combinées, leur différence brute étudiée, le tout sans équivoque. Prenons un exemple du contraire : on veut comparer (ou combiner) la statistique d'un groupe de $n = 10$ éléments ($p_{10} = x_{10}/10$) à celle d'un groupe de $n = 200$ ($p_{200} = x_{200}/200$). Le premier obtient $p_{10} = 1/10 = 0,100$, le second, $p_{200} = 25/200 = 0,125$: le second groupe a donc un score vraiment plus élevé que le premier. Selon la fonction y_{ANS} d'Anscombe, nous calculons $y_{1:10} = 0,366$ et $y_{25:200} = 0,363$, ce groupe apparaissant maintenant égal au premier, sinon légèrement plus faible. Si maintenant le second groupe obtenait un x de 20, égalisant la proportion donnée par le premier groupe, nous aurions alors $y_{20:200} = 0,324$, une valeur carrément plus basse que le 0,366 du premier groupe alors que leurs proportions brutes sont en fait égales. La fonction y_{CHA} de Chanter, qui corrige pour le biais, rétablit, mais en partie seulement, la comparabilité des proportions basées sur des tailles n inégales.

Rappelons que le biais attesté de la fonction angulaire est structurel (ou mathématique) et non statistique, et il n'est remédié que par l'accroissement de la taille n sur laquelle est basée la proportion.¹ C'est ce double biais, selon π et n , que nous examinons maintenant.

Les graphiques B1 à B6 (placés en fin de texte) illustrent l'évolution du biais des fonctions y_{ANS} , $y_{F-T/2}$ et y_{CHA} selon n , soit l'espérance de chacune selon π (le paramètre binomial) et n , en fonction de n : seules quelques valeurs dans l'intervalle $0 < \pi \leq 1/2$ sont présentées, la fonction (avec son biais) étant antisymétrique selon π . Les fonctions approchent de l'asymptote (égale à $\sin^{-1}\sqrt{\pi}$) dès $n = 25$ et s'y stabilisent à peu près à $n = 100$. La fonction y_{CHA} accuse un net avantage sur les autres, étant déjà prati-

1. La variation au hasard de la variable binomiale $x \sim B(\pi, n)$ est asymétrique dès que $\pi \neq 1/2$ et crée elle-même un biais statistique, lequel s'atténue avec l'asymétrie à mesure que n croît.

quement sans biais à $n = 25$, la fonction y_{ANS} affichant par ailleurs un biais très important dans les cas de petites tailles n . La figure B6 illustre à l'évidence l'absence de tout biais sous $\pi = 1/2$.

Dans les graphiques suivants, de B7 à B11, nous avons tenté de montrer ce qu'il en est de l'estimation de la proportion p , c.-à-d. de l'espérance de cette proportion) après qu'elle soit passée par chaque fonction de transformation : est-ce que la p résultante approche ou non de la valeur du paramètre π qui l'a nourrie. L'estimation brute est obtenue en inversant la fonction asymptotique, soit $(\sin[\epsilon(y)])^2$; une estimation plus serrée consiste à extraire du résultat l'estimation du numérateur x correspondant, puis en refaisant le quotient x/n . Cette reconstitution de l'estimateur $\hat{\pi}(n)$ décline nettement la fonction $F-T$ et donne un léger avantage à l'estimateur de Anscombe sur celui de Chanter. Le lecteur aura noté une sous-estimation systématique de π pour le demi-domaine $0 < \pi < 1/2$ à quoi répond une surestimation en miroir pour le domaine $1/2 < \pi < 1$.

La fonction y_{CHA} de CHANTER (1975) gagne sans grand mérite le concours du biais le plus acceptable (notamment dans la métrique de la transformation angulaire, celle avec laquelle les tests statistiques sont à effectuer) puisque Chanter l'a expressément façonnée à cette fin. Les deux autres fonctions se font la lutte dans les petites valeurs de n , avec un mince avantage de y_{ANS} contre $y_{F-T/2}$. Il reste que, selon nous, le biais est une propriété importante, voire la plus importante, d'une fonction de calcul, de sorte que, toutes choses étant égales d'ailleurs, la primauté de y_{CHA} doit être considérée. D'un autre côté, rappelons que la présence de biais structurel n'a en soi nulle incidence sur le traitement (p. ex. la comparaison) de proportions basées toutes sur la même taille n .

La variance des fonctions

Les trois fonctions proposées de la transformation angulaire (incluant la fonction modifiée $y_{F-T/2}$) ont, asymptotiquement sur n , une variance standardisée de $n \times \sigma^2(y) = 1/4$. Toutefois, la valeur réelle de cette variance dépend aussi de n et, bien sûr, de π . Les graphiques V1 à V5 font voir le chemin selon n de l'espérance $n \times \sigma^2(y)$ des trois fonctions vers l'asymptote $1/4$, pour quelques valeurs représentatives de π entre 0 et $1/2$ (la variation pour π de $1/2$ à 1 est symétrique). La très faible variance de la source binomiale pour $\pi \rightarrow 0$ ou $\pi \rightarrow 1$ rend difficile l'étalement désiré par la transformation angulaire. Par exemple, avec $n = 100$ et $\pi = 0,01$, la variance binomiale de $n \times \pi \times (1 - \pi) \approx 1$ ne laisse guère de manœuvre pour l'étalement, et cela devient de plus en plus impossible pour de plus petits n .

Même si les trois fonctions montrent des allures similaires, on note que y_{ANS} monte plus paresseusement

et plus tardivement vers la cible alors que, talonnée par $y_{F-T/2}$ dans l'intervalle $0 < \pi < 0,10$, y_{CHA} tend à dominer ensuite et colle davantage à la cible. À toutes fins pratiques, l'approche des variances vers $1/4$ est satisfaisante pour $\pi \approx 0,05$ dès $n = 25$, pour $\pi \approx 0,20$ dès $n = 5$. CHANTER (1975) présente aussi, dans sa Figure 1, des graphiques de la variance des fonctions y_{ANS} et y_{CHA} , de même que la fonction simple, $\sin^{-1}\sqrt{x/n}$, cette dernière affichant un comportement plutôt excentrique selon n . Quant à y_{ANS} et y_{CHA} , l'élévation de leurs variances pour rejoindre l'asymptote $1/4$ est semblable, avec un léger avantage précoce pour y_{CHA} .

La précision des fonctions via les Écarts Quadratiques Moyens

L'écart (ou erreur) quadratique moyen (EQM) mesure la marge de variation typique d'une fonction statistique par rapport à sa valeur cible, ici l'écart entre chaque fonction « y » et la cible $\sin^{-1}\sqrt{\pi}$, selon $EQM = \epsilon(y - \sin^{-1}\sqrt{\pi})^2$. L' EQM dépend à la fois de la variance de y et de son biais, selon l'équation $EQM \approx \sigma^2(y) + [\epsilon(y) - \sin^{-1}\sqrt{\pi}]^2 = \text{Variance} + \text{Biais}^2$. Les graphiques EQM1 à EQM8 montrent une marche grossièrement comparable des trois fonctions vers la précision à mesure que n croît, le biais et la variance de chacune s'amointrissant. L'avantage passe d'une fonction à l'autre, selon π . Pour π de 0,01 à 0,025, les fonctions y_{Cha} et $y_{F-T/2}$ dominent, passant à y_{Cha} et y_{Ans} pour π de 0,05 à 0,20, puis à y_{Ans} pour 0,30 à 0,40 et enfin à y_{Ans} et $y_{F-T/2}$ pour 0,50, les inégalités restant toutefois mineures.

Les EQM standardisés, affichés aux graphiques EQM9 à EQM16, sont simplement le produit $n \times EQM$ et ils compensent pour l'effet de n sur la part de variance qui y est incluse : la cible de cette statistique est ici $1/4$, soit la variance paramétrique des fonctions « y », que rejoint d'abord $y_{F-T/2}$ pour π de 0,01 à 0,05, puis y_{Cha} et $y_{F-T/2}$ pour 0,10 à 0,20 et que colle enfin y_{Cha} pour 0,30 à 0,50. La stabilisation de la quasi double fonction $y_{F-T/2}$ y est pour quelque chose, comme l'est pour y_{Cha} la correction expresse du biais pour n . En termes de précision, les trois fonctions considérées sont acceptables, avec un léger bémol pour la fonction d'Anscombe.

La normalité des fonctions et leurs indices de forme γ_1 et γ_2

L'indice d'asymétrie $\gamma_1 = \epsilon(X - \mu)^3/\sigma^3$ et l'indice de voussure ou aplatissement $\gamma_2 = \epsilon(X - \mu)^4/\sigma^4 - 3$ tous deux égaux à 0 dans une distribution normale, servent souvent d'indicateurs de la normalité approximative d'une statistique; les indices statistiques correspondants, calculés comme de simples moyennes sur plusieurs échantillons, sont notés g_1 et g_2 . Les fonctions y_{Ans} , y_{Cha}

Tableau 1 ■ Valeur minimale de π requise pour obtenir $N = |\gamma_1| + 1/2 |\gamma_2| < 0,5$ selon n

n	π minimal	Anscombe	Chanter	F-T
5	0,31	0,22	0,22	0,20
10	0,2	0,12	0,12	0,11
25	0,13	0,05	0,06	0,07
50	0,09	0,03	0,03	0,04
100	0,05	0,02	0,02	0,02
500	0,02	0,01	0,01	0,02
1000	0,01	0,01	0,01	0,01

Note. Chaque inscription π du tableau peut être interprétée $1 - \pi$.

et $y_{F-T/2}$ sont-elles normales ?

Les graphiques N1 à N7 étalent les valeurs de γ_1 et γ_2 (à calcul exact) de nos trois fonctions ainsi que de la proportion brute, ce pour des tailles d'échantillon $n = 5, 10, 25, 50, 100, 500$ et 1000 , selon π variant de $0,05$ à $0,50$. Globalement, les fonctions « y » comme la proportion simple présentent une asymétrie positive décroissante de $\pi = 0^+$ à $0,50^-$ et négative croissante de $\pi = 0,50^+$ à 1^- , l'une étant en miroir de l'autre. Quant à la voussure, son dessin selon π est convexe : symétriquement positif (léptokurtique ou pointu) aux valeurs extrêmes de π et descendant vers un creux plus ou moins bossué (mésokurtique ou normal si $\gamma_2 \approx 0$ et platykurtique, aplati si $\gamma_2 < 0$) vers le centre, les fonctions « y » présentant un certain redressement central. Les graphiques ce celles-ci montrent un comportement généralement plus « normal », à indices moins loin de 0, que la proportion simple.

Quand peut-on considérer que la variable est suffisamment « normale », c.-à-d. acceptable comme ingrédient légitime d'un test statistique (t de Student, analyse de variance, régression linéaire, etc.) ? L'indice γ_1 étant vraisemblablement plus critique que γ_2 , on note qu'une distribution à valeurs sous $\gamma_1 \approx 0,20$ (p. ex. $\gamma_1 \approx 0,17$ pour y_{ANS} à $n = 5$ et $\pi = 0,25$) présente une asymétrie visuellement indécélable, et qu'elle reste très modérée à $\gamma_1 = 0,50$: de toutes façons, les trois fonctions « y » ont toutes et presque partout des valeurs de non-normalité plus faibles que la proportion brute. Le tableau 1 peut guider le lecteur : il s'agit d'un indice (N) de non-normalité composée, construit en combinant γ_1 et γ_2 (en valeurs absolues), pondérés par leurs variances approximatives respectives.

Pour chaque taille n présentée, le tableau 1 indique la valeur π minimale pour obtenir un indice $N \leq 0,50$, une valeur relativement sévère (parce que combinant deux indices). On constate que, dès $\pi \approx 0,20$ ou $0,22$ (ou sous $0,78$ à $0,80$), les trois fonctions ont un comportement déjà acceptable à $n = 5$, et que les tailles n de 50 et plus rendent abordables des π assez extrêmes ($\leq 0,03$ ou $\geq 0,97$).

CHANTER (1975), dans son Tableau 2, présente son étude

sur la non-normalité de diverses fonctions, notamment p (la proportion brute), y_{ANS} et y_{CHA} : voir notre tableau 2.

La transformation angulaire rapproche indéniablement la variable d'une répartition normale, de façon quasi complète à $n = 100$ pour tout π et dès $n = 25$ pour $0,02 \leq \pi \leq 0,95$.

L'exactitude des fonctions (ou le respect du seuil de significativité α)

Le test statistique de presque n'importe quelle variable continue est capable de rencontrer parfaitement le seuil de significativité exact ordinairement exigé d'une variable normale. Par exemple, la comparaison des moyennes de deux échantillons de distributions respectives $f(\mu_1, \Theta)$ et $f(\mu_2, \Theta)$, où Θ dénote ici un groupe de paramètres et μ l'espérance de la variable, peut être amenée à respecter tout seuil de significativité α grâce à une transformation de variable appropriée, qu'elle se réfère ou non à la loi normale. Dans le cas de variables *discrètes*, comme ici les proportions (les numérateurs sont des nombre naturels), non seulement la cible α ne peut qu'être approximée mais elle peut aussi être difficile à approcher : les deux obstacles sont la granularité grossière de la variable sous de petites tailles n (le numérateur étant confiné dans l'ensemble de valeurs $[1, 2, \dots, n]$, ce qui transparait dans le caractère *saltatoire* de la distribution) et, pour la loi binomiale et d'autres lois semblables, l'asymétrie de la distribution de probabilités, laquelle devient difficilement contrôlable lorsque $\pi \rightarrow 0$ ou $\pi \rightarrow 1$.

Cette difficulté particulière de respecter l'exactitude du test de significativité sur des proportions s'applique a fortiori sur un test qui serait de calcul binomial exact (tel que le calcul de la probabilité extrême associée à une proportion simple ou les tests de LIDDELL (1978) et de LAURENCELLE (2005, 2017, 2021b) pour une différence de proportions. Ces tests respectent, c.-à-d. se basent directement sur la discontinuité de la variable, d'où le caractère saltatoire de la probabilité associée, discontinuité qui affecte aussi les fonctions normalisantes telles que celles explorées ici. Les tests par ces fonctions donnent-ils

Tableau 2 ■ Proportion du domaine de π pour laquelle les fonctions étudiées présentent des indices γ_1 et γ_2 pas trop éloignés de la normalité (données reprises de Chanter 1975, p. 358)

	$n = 10$		$n = 100$	
	$ \gamma_1 < 1/2$	$ \gamma_2 < 1$	$ \gamma_1 < 1/2$	$ \gamma_2 < 1$
$p = x/n$	0,621	0,867	0,929	0,981
y_{ANS}	0,846	0,945	0,985	0,995
y_{CHA}	0,855	0,948	0,986	0,996

une sanction fiable de la significativité d'une différence, la fiabilité dépendant essentiellement du respect du seuil de significativité? C'est ce que nous avons voulu illustrer par des études spécifiques (calculs binomiaux exacts ou échantillons de type Monte Carlo basés sur 100 000 échantillons ou plus), études dans lesquelles nous plaçons en lice nos fonctions y_{ANS} , y_{CHA} et $y_{F-T/2}$. Pour le test sur une proportion, la précision est évaluée par le critère du calcul binomial exact; le test z avec correction de $\pm 1/2$ est aussi inclus pour comparaison. Quant au test de différence entre deux proportions, le test classique $z(p_1 - p_2)$ sans correction de continuité² et le calcul binomial exact de LAURENCELLE (2017, 2021b)³ sont aussi considérés.

Le test sur une proportion

Les graphiques Ex1 à Ex8 font voir les résultats d'une courte étude (à calculs exacts) de l'aptitude des fonctions « y » à reproduire précisément la décision de significativité donnée par le calcul binomial, ce au seuil de significativité unilatéral droit de $\alpha = 0,05$.

Pour les illustrations Ex1 à Ex4, la valeur x^* tout juste significative [i.e. $p^* = p(x^*|n, \pi) \leq \alpha \leq p(x^* + 1|n, \pi)$] est d'abord trouvée, puis la probabilité 'normale' $p(z|x^*)$ associée à chaque test z (Brute, y_{Ans} , y_{Cha} , $y_{F-T/2}$) pour cette valeur x^* est obtenue, et enfin la différence $d = p(z|x^*) - p^*$ est calculée. Le graphique Ex1 montre la différence moyenne $\bar{d}$ calculée pour des tailles n de 5 à 20 se rapprochant de 0 selon quatre niveaux du paramètre π , correspondant à une variable x très asymétrique sous $\pi = 0,1$ vers une symétrique sous $\pi = 0,5$. Les fonctions angulaires surestiment toutes la probabilité exacte p^* , la plus proche étant $y_{F-T/2}$ et la plus lointaine y_{Ans} : cette surestimation est susceptible d'enlever un peu de sensibilité de ces fonctions à une différence réelle. La fonction 'Brute', soit ici $z_{Brute}(x^*) = (x^* - \frac{1}{2} - n \cdot \pi) / \sqrt{n \cdot \pi \cdot (1 - \pi)}$, quant à elle, sous-estime la probabilité p^* , ce qui peut au contraire aboutir à déclarer signifi-

cative une différence qui ne l'est pas. Tel qu'attendu, l'ampleur de ces mésestimations décroît des graphiques Ex2 à Ex4, à mesure que la taille des échantillons croît, les distributions de la variable x se rapprochant alors de la forme normale.

Jusqu'à quel point les sanctions de significativité des fonctions angulaires tombent-elles juste, c'est ce à quoi répondent les graphiques Ex5 à Ex8. Ils affichent le taux de concordance moyen des fonctions angulaires avec le calcul exact, c.-à-d. quant au fait que la valeur x' de chaque fonction tout juste significative précède, égale ou excède la valeur x^* exacte : le seuil de signification α appliqué est de 0,05. En fait, les quatre graphiques ne rapportent que la concordance précise, celle où $x' = x^*$, les autres cas lui étant très redondants. Trois remarques s'imposent sur ces données : (1) La fonction $y_{F-T/2}$ donne la performance la plus juste des quatre tests dans le cas de variables asymétriques ($\pi \neq 1/2$); (2) La proportion de non-concordance ($x' \neq x^*$), complémentaire à la proportion affichée dans les graphiques, est de deux types : dans le cas des fonctions angulaires y_{Ans} , y_{Cha} et $y_{F-T/2}$, la non-concordance est toujours de type ($x' > x^*$), c.-à-d. que le test n'y ressort significatif qu'à la valeur $x^* + 1$, alors qu'au contraire, pour le test classique 'Brute', il l'est toujours à $x^* - 1$ ou avant, donc faussement significatif; (3) Des quatre tests comparés, l'ordre de justesse se décline comme $y_{F-T} > y_{Ans} > y_{Cha} > Brute$, sauf dans le cas symétrique à $\pi = 1/2$ pour lequel y_{Cha} et Brute dominent, confirmant encore une fois l'excellence de l'approximation normale (par le z_{Brute}) dans le cas d'une binomiale symétrique.

Les résultats affichés reflètent le manque de continuité numérique (de probabilité, de présence de fonction, de niveau de significativité) toujours souhaitable pour un test de significativité statistique, un effet certain de la discontinuité de la variable binomiale x . Le tableau 3 peut aider à mieux appréhender l'importance et le risque d'im-

2. La correction de continuité, légitimement appliquée au test d'une proportion, est ici rendue inopérante par le calcul de différence entre deux proportions, les deux corrections s'annulant par la soustraction au numérateur (exactement si $n_1 = n_2$ ou approximativement sinon).

3. Le test à calcul exact de LIDDELL (1978), basé sur une estimation ponctuelle de la probabilité sous H_0 (soit l'estimation de π par maximum de vraisemblance) donne des sanctions pratiquement identiques à celles de notre calcul plus complet (voir LAURENCELLE, 2017, 2021b), et les probabilités produites pour chaque échantillon Monte Carlo sont très proches l'une de l'autre ($r \geq 0,9995$). Cependant, sous H_0 et π constant, la variance du calcul Liddell tend à être moindre pour π proche de 0 ou de 1 que celle de Laurencelle, et les probabilités produites sont alors différentes.

Tableau 3 ■ Saut de continuité autour de la valeur critique (x^*) binomiale sous $B(\pi = 0, 3, n)$

n	x^*	$Pr(x \geq x^*)$	$x^* - 1$	$Pr(x \geq x^* - 1)$
hline 10	6	0,0473	5	0,1027
25	12	0,0442	11	0,0978
50	21	0,0478	20	0,0848
100	39	0,0340	38	0,0530
500	168	0,0449	167	0,0547

FIGURE 1 ■ Probabilités binomiales $B(\pi = 0, 1, n = 100)$ vs $B(\pi = 0, 1, n = 10)$

pact de la discontinuité de la variable binomiale. Il s'agit ici d'une binomiale $B(\pi = 0, 3, n)$, où se trouvent placées côte à côte la valeur x^* tout juste significative telle que $Pr\{x \leq x^*\} \leq 0,05$ et la valeur immédiate inférieure ($x = x^* - 1$) : le lecteur peut le constater, les écarts de continuité peuvent être importants, et ils s'appliquent à la variable x d'origine comme aux transformations qu'on lui impose.

En conclusion, discontinue certes, la probabilité estimée par les fonctions « y » est discontinue, saltatoire, comme l'est celle obtenue par un calcul binomial exact, et toutes respectent, avec un écart de probabilité plus ou moins important, le seuil α prescrit tout en concordant généralement avec la décision binomiale. Le test binomial exact, ici de calcul facile, ne requiert pas de table de valeurs critiques spécifiques pour chaque couple π et n , et c'est lui que nous recommandons. Enfin, les fonctions angulaires « y » (de type $\sin^{-1} \sqrt{y(x, n)}$) émulent assez bien, voire dominent généralement l'approximation normale classique.

Le test de différence entre deux proportions

Nous nous sommes intéressé au test de différence entre deux proportions pour son propre mérite et, aussi et sur-

tout, parce qu'il concerne un composé linéaire de k variables où $k = 2$, donnant un aperçu du comportement des fonctions y pour des composés ayant $k \geq 2$ ingrédients tels qu'on en rencontre en analyse de variance.

Nous l'avons déjà signalé, la discontinuité de la variable binomiale et sa forte asymétrie lorsque son paramètre π s'écarte de $1/2$ ont un effet important sur l'exactitude des tests, un effet d'autant plus important lorsque deux binomiales à configurations distinctes sont comparées. En voici un exemple, soit une $B(\pi = 0, 1, n = 10)$ et une $B(\pi = 0, 1, n = 100)$. La figure 1 fait voir les tracés de leurs probabilités.

L'asymétrie positive de $B(10, \pi)$ déplace sa médiane vers la droite,⁴ bornant des valeurs de x plus élevées, de sorte que le test unilatéral droit $B(10, \pi) \geq B(100, \pi)$ est avantage par rapport au test $B(100, \pi) \geq B(10, \pi)$. Autrement dit, la variance réduite et la symétrie quasi normale de $B(100, \pi)$ placent ses valeurs près de l'espérance $\sim \pi = 0, 1$, ce qui les prévient d'excéder les valeurs de $B(10, \pi)$ étalées à droite. Ces configurations de probabilités donnent lieu aux résultats de tests de la figure 2.

Comme on voit, sous H_0 au seuil $\alpha = 0,05$, les différences $p(B(100; 0, 1) - p(B(10; 0, 1))$ à la figure de gauche réussissent à peine à s'élever vers le seuil α pres-

4. Les proportions médianes interpolées sont de 0,391 et 0,369 pour les distributions $B(0, 1; 10)$ et $B(0, 1; 100)$ respectivement.

FIGURE 2 ■ Test sur deux proportions pour $\pi = 0, 1$ (gauche) et $\pi = 0, 3$ (droite)

crit, au contraire des différences opposées $p(B(10; 0, 1) - p(B(100; 0, 1))$, qui s'y rendent joyeusement. Le lecteur aura noté que cet effet, pour le moins paradoxal, affecte aussi bien le calcul dit exact que les approximations, notamment l'approximation « Brute » classique, soit $z(p_1 - p_2) = (p_1 - p_2)/\sqrt{P(1-P)(1/n_1 + 1/n_2)}$, où $P = (n_1 p_1 + n_2 p_2)/(n_1 + n_2)$. Cet effet s'estompe dès avant $\pi \approx 0, 3$, tel qu'illustré à la figure de droite, et il disparaît à $\pi = 0, 5$ pour réapparaître, inversé cette fois, pour $\pi > 0, 5$ et $\pi \rightarrow 1$.

Les 35 graphiques annexés font voir les effets d'une panoplie de conditions sur le degré d'exactitude des cinq tests considérés. Les 10 premiers graphiques (Ex9 à Ex18) traitent de groupes à tailles égales, ce pour les seuils $\alpha = 0, 05$ et $0, 01$, tandis que les 20 suivants (Ex19 à Ex38) explorent deux formes d'inégalités; les 5 derniers graphiques, Ex39 à Ex43, fournissent une démonstration éloquentes des effets encourus par ces inégalités de tailles. Le test appliqué est toujours de mode unilatéral, à droite.

Les graphiques présentés font voir une imprécision plus ou moins grande des fonctions « y » - et aussi du calcul binomial exact par rapport au respect du seuil α imposé. Ces graphiques ont pour première fonction de faire ressortir les disparités ou errements possibles des données obtenues, une propriété qui peut avoir l'effet pervers de rendre le lecteur suspicieux à l'égard de la fiabilité des statistiques étudiées. La prudence, oui, est toujours de bon aloi, ici comme ailleurs. Cependant le lecteur doit considérer (1) que les variations de probabilité autour de la valeur α prescrite sont, dans la plupart des conditions, assez restreintes en chiffres absolus et (2) que, à l'instar d'autres fonctions de calcul, il s'agit ici d'estimateurs échantillonnaux ayant chacun son erreur-type et traitant avec des valeurs ob-

servées à charge aléatoire; l'effet massif de la discontinuité de la variable étudiée (voir plus haut) doit aussi être pris en compte.

L'imprécision visible de la sanction du test de différence et son observance approximative du seuil α prescrit sont, on le voit, causées principalement par la nature même de la variable, et elles affectent aussi bien sinon davantage le « calcul exact » et le test $z(p_1 - p_2)$ classique. L'inégalité des tailles (n) y rajoute aussi, effet déconcertant (puisque'il est logiquement imprévisible) que nous assignons à la confrontation de deux asymétries différentes, tel qu'illustré en début de section. Ces effets sont caricaturés par les dernières figures, aux numéros Ex39 à Ex43. L'ensemble des graphiques concernant le test de différence de deux proportions affiche une similitude attendue de nos trois fonctions angulaires avec, occasionnellement, un écart parfois avantageux mais plus souvent défavorable à la fonction $y_{FT/2}$, les deux autres (y_{Ans} et y_{Cha}) demeurant bonnes voisines.

Un effet qu'il nous a paru intéressant de constater est que, si d'une part la probabilité du test exact se tient toujours nettement et évidemment sous le seuil α prescrit, d'autre part la probabilité des tests approximatifs frôle souvent et semble fluctuer près du seuil, pouvant donner l'illusion que ces tests sont parfois plus « exacts » que le calcul binomial. Cette illusion, qui n'est pas déplaisante, a tout de même de quoi nous rassurer sur la pertinence des fonctions angulaires pour l'étude de la différence entre deux ou plusieurs proportions.

Enfin, même si les mathématiques dans lesquelles s'exprime le discours statistique sont tout sauf imprécises, le discours lui-même, son contenu, reste estimatif et sa conclusion, toujours incertaine. L'avis de l'auteur est que

mieux valent des estimateurs statistiques un peu imprécis (mais efficaces, comme la section suivante le démontrera) qu'une simple divination basée sur le « bon sens » : le chercheur n'est pas un aruspice. Le lecteur sera peut-être aussi intéressé à l'étude de LAURENCELLE (2017), dans laquelle le test binomial exact est mis en compétition avec d'autres tests, plus classiques ceux-là, de la différence entre deux proportions.

La puissance des fonctions

Nous n'avons étudié la puissance statistique associée aux fonctions « y » sous enquête qu'à travers le test de différence entre deux proportions (indépendantes), ce en mode unilatéral à droite, soit explicitement $Pu = Pr\{y(\pi_1, n_1) \geq y(\pi_2, n_2), \alpha\}$, selon $\pi_1 \geq \pi_2$. L'étude de la puissance du test sur une proportion versus une référence paramétrique reste à faire. Le lecteur intéressé à la question peut aussi recourir à d'autres approches d'étude que celles explorées ici, notamment l'excellente approximation normale Gram-Charlier pour une proportion (LAURENCELLE & DUPUIS, 2002b, p. 184; ou LAURENCELLE & DUPUIS, 2002a, p. 166; KENDALL & STUART, 1977). En outre, l'intérêt particulier des fonctions dites angulaires tient à leur combinabilité linéaire, c.-à-d. l'aptitude qu'elles offrent à être additivement composées grâce à leur variance stabilisée et leur quasi normalité, le calcul d'une moyenne de quelques valeurs, le test de différence entre deux ou plusieurs valeurs, la validité des techniques de traitement à base de normalité devenant tous possibles.

Les 27 graphiques rapportés couvrent différentes conditions paramétriques, notamment les tailles n_1 et n_2 égales (Pu1 à Pu9) et inégales (Pu10 à Pu27), ces dernières comparant les inégalités $n_1 < n_2$ à $n_1 > n_2$, c.-à-d. les cas où les proportions « plus fortes » (sous π_1) émanent d'un échantillon plus petit ($n_1 < n_2$) ou plus grand ($n_1 > n_2$). Tous les cas rapportés appliquent le seuil α de 0,05; nos données non rapportées pour le seuil de 0,01 sont d'effets semblables.

Pour les comparaisons à tailles échantillonnales égales, les quatre fonctions « y » coïncident presque partout, montrant un léger avantage sur le calcul exact (décrit par la ligne pointillée de chaque figure) dès que les tailles n débordent 5 et 10, toutes se rejoignant à $n = 50$. Du côté des tailles n_1 et n_2 inégales, nous n'avons pas trouvé d'avantage net entre la condition $n_1 > n_2$, dans laquelle la proportion à paramètre π plus élevé est aussi la plus précise (et mieux empiriquement définie) et la condition opposée, $n_1 < n_2$. Certains effets curieux apparaissent néanmoins, par exemple entre les conditions 10 vs 50 et 60 vs 10 ou 10 vs 100 et 100 vs 10, pour lesquelles, anecdotiquement sous $\pi = 0, 1$, la fonction y_{Brute} passe du premier au dernier rang, tandis que, pour les tests à petites

tailles ($n = 5$ et 10), la fonction $y_{F-T/2}$ montre un net mais éphémère avantage sur ses concurrentes.

Noter en dernier lieu que, à notre connaissance, il n'existe aucune procédure de calcul binomial raisonnablement commode pour la comparaison de $k > 2$ proportions indépendantes.

Propos conclusifs

L'étude faite des propriétés des fonctions étudiées et utilisées à des fins d'analyse statistique, soit leur distribution, leur avantage sur l'utilisation de la proportion brute, leur validité pour effectuer des tests d'hypothèses statistiques, n'est évidemment ni exhaustive ni rigoureuse. Nous nous risquons malgré tout à une conclusion globale, en deux volets. (1) Lorsqu'il est possible, le calcul exact, soit celui de la loi binomiale simple pour le test sur une proportion, ou celui de LAURENCELLE (2017, 2021b) ou son approximation par LIDDELL (1978) pour le test sur deux proportions, reste... exact et serait idéalement à préférer. (2) Les 4 fonctions étudiées, particulièrement les trois recourant à la transformation angulaire inspirée de Fisher, émulent généralement bien le calcul exact. De plus, leur variance moins dépendante de l'espérance que dans le cas de la proportion p et leurs propriétés linéaires plus proches de la variable normale font qu'il est concevable de les exploiter prudemment à titre de statistiques sommatives, ou moyennes, à variance stabilisée, et de les intégrer dans des procédures telles que l'analyse de variance.

Références

- ANSCOMBE, F. J. (1948). The transformation of Poisson, binomial and negative-binomial data. *Biometrika*, 35, 246-254.
- CHANTER, D. O. (1975). Modifications of the angular transformation. *Journal of the Royal Statistical Society Series B (Applied Statistics)*, 24, 354-359.
- DEVROYE, L. (1986). *Non-uniform random variate generation*. New York : Springer-Verlag.
- FREEMAN, M. F. & TUKEY, J. W. (1950). Transformations related to the angular and the square root. *Annals of Mathematical Statistics*, 21, 607-611.
- KENDALL, M. & STUART, A. (1977). *The advanced theory of statistics. Vol. 1 : Distribution theory*. London : Griffin.
- LAURENCELLE, L. (2001). *Hasard, nombres aléatoires et méthode Monte Carlo*. Québec : Presses de l'Université du Québec.
- LAURENCELLE, L. (2005). Le test binomial exact de la différence entre deux proportions. *Lettres statistiques*, 12, 1-13.
- LAURENCELLE, L. (2017). Le test binomial exact de la différence entre deux proportions et ses approximations [The exact binomial test between two propor-

tions and its approximations]. *The Quantitative Methods for Psychology*, 13, 141-147.

LAURENCELLE, L. (2021b). The exact binomial test between two independent proportions : A companion. *The Quantitative Methods for Psychology*, 17(2), 76-79. doi :[10.20982/tqmp.17.2.p076](https://doi.org/10.20982/tqmp.17.2.p076)

LAURENCELLE, L. & DUPUIS, F. (2002a). *Statistical tables, explained and applied*. Singapour : World Scientific.

LAURENCELLE, L. & DUPUIS, F. (2002b). *Tables statistiques expliquées et appliquées*. Québec : Le Griffon d'argile.

LIDDELL, D. (1978). Practical tests of 2 x 2 contingency tables. *The Statistician*, 25, 295-304.

RÜCKER, G., SCHWARZER, G., CARPENTER, J. & OLKIN, I. (2009). Why add anything to nothing? The arcsine difference as a measure of treatment effect in meta-analysis with zero cells. *Statistics in Medicine*, 28, 721-738.

WANG, K. & FINGERT, J. H. (2012). Statistical tests for detecting rare variants using variance stabilizing transformations. *Annals of Human Genetics*, 76, 402-409.

Citation

LAURENCELLE, L. (2021a). Étude comparative de trois fonctions de transformation d'une proportion (Anscombe, 1948 – Freeman-Tukey, 1950 – Chanter 1975). *The Quantitative Methods for Psychology*, 17(3), 166-194. doi :[10.20982/tqmp.17.3.p166](https://doi.org/10.20982/tqmp.17.3.p166)

Copyright © 2021, LAURENCELLE. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 08/01/2021 ~ Accepted: 21/05/2021

Les figures B1 à Pu27 suivent.

Figures B (partie i) ■ Biais. Les graphiques B1 à B6 illustrent le biais subasymptotique des fonctions angulaires. Ce biais sur π est antisymétrique, les fonctions surestimant l'espérance vers $\pi = 0$ et la sous-estimant vers $\pi = 1$.

(B1)

(B2)

(B3)

(B4)

(B5)

(B6)

Figures B (partie ii) ■ Biais (suite). Les graphiques B7 à B11 illustrent le biais la forme d'une reconstitution de la proportion estimée π' à partir de l'espérance calculée pour chaque fonction angulaire y . La fonction est inversée selon $p'(y) = [\sin(\epsilon(y(x, n)))]^2$, la valeur correspondante du numérateur x est calculée (p. ex., pour la fonction Anscombe, $x' = p' \times (n + 3/4) - 3/8$, puis $\pi'(n) = x'/n$. La droite noire indiquée « $Esp(\infty)$ » représente la limite asymptotique $\pi'(\infty)$ de la statistique.

(B7)

(B8)

(B9)

(B10)

(B11)

Figures V ■ Variance. Les graphiques V1 à V6 montrent les variances standardisées ($n \times \sigma^2(y|n, \pi)$) présentées sont obtenues par un calcul binomial exact. Noter que les courbes d'évolution pour $1/2 \leq \pi < 1$ sont symétriques à celles données ici.

(V1)

(V2)

(V3)

(V4)

(V5)

(V6)

Figures EQM (partie i) ■ EQM (ou précision). Les graphiques EQM1 à EQM16 montrent les écarts quadratiques moyens (*EQM*), obtenus pour chaque fonction y par $\epsilon\{[(y|n, \pi) - \sin^{-1}\sqrt{\pi}]^2\}$, reflètent l'imprécision de la statistique y due à la fois à la variabilité naturelle et au biais de la fonction utilisée, les deux étant amortis par la taille n . Les *EQM* « standardisées (= $n \times EQM$) » permettent une comparaison relative du comportement des fonctions sur une échelle homogène, c.-à-d. à variance d'erreur à peu près égale.

(EQM1)

(EQM2)

(EQM3)

(EQM4)

(EQM5)

(EQM6)

Figures EQM (partie ii) ■ EQM (ou précision, suite)

(EQM7)

(EQM8)

(EQM9)

(EQM10)

(EQM11)

(EQM12)

Figures EQM (partie iii) ■ EQM (ou précision, suite et fin)

(EQM13)

(EQM14)

(EQM15)

(EQM16)

Figures N (partie i) ■ Normalité. Les graphiques N1a à N7b montrent les indices γ_1 et γ_2 qui reflètent respectivement le degré d'asymétrie et le degré d'aplatissement (ou voussure) de la distribution d'une statistique, la distribution normale ayant $\gamma_1 = 0$ et $\gamma_2 = 0$. Les valeurs représentées aux graphiques sont de calcul exact. Pour la statistique de proportion « brute », $p = x/n$, $\gamma_1 = (1 - 2\pi)/\sigma$ et $\gamma_2 = [1 - 6\pi(1 - \pi)]/\sigma^2$, où $\sigma^2 = \pi(1 - \pi)/n$.

(N1)

(N2)

(N3)

(N4)

(N5)

(N6)

Figures N (partie ii) ■ Normalité (suite).

(N7)

(N8)

(N9)

(N10)

(N11)

(N12)

Figures N (partie iii) ■ Normalité (suite et fin).

(N13)

(N14)

Figures Ex (partie i) ■ Exactitude - Tests sur 1 proportion (seuil $\alpha = 0,05$). Les graphiques Ex1 à Ex4 présentent, pour quatre intervalles de tailles n croissantes, la différence de probabilité $p_y(x^*) - p_{Bin}(x^*)$ de chaque test approximatif « y » par rapport au calcul binomial exact tel que $p_{Bin}(x^*) \leq \alpha \leq p_{Bin}(x^* + 1)$, ce en fonction de différentes valeurs de π . Quant aux graphiques Ex5 à Ex8, pour quatre valeurs de π , ils indiquent le taux de concordance entre les valeurs x^* (tel que définie ci-dessus) et x' , où x' est obtenue par chaque test « y » comparé telle que $p_y(x') \leq \alpha < p_y(x' + 1)$, la concordance étant définie selon $x' = x^*$.

(Ex1)

(Ex2)

(Ex3)

(Ex4)

Figures Ex (partie ii) ■ Exactitude - Tests sur 2 proportions - Les graphiques Ex5 à Ex8 documentent des comparaisons pour tailles n_1 et n_2 égales (seuil $\alpha = 0,05$). La donnée d'exactitude présentée est la « puissance » moyenne de chaque test (500 000 échantillons) pour atteindre le seuil fixé. Certains graphiques présentent des lignes superposées, réunissant des fonctions « y » sous la couleur de la dernière tracée (souvent le « jaune » de la fonction $y_{F-T/2}$).

(Ex5)

(Ex6)

(Ex7)

(Ex8)

Figures Ex (partie iii) ■ Graphiques - Tests sur 2 proportions - Tailles égales (seuil $\alpha = 0,01$)

(Ex9)

(Ex10)

(Ex11)

(Ex12)

(Ex13)

Figures Ex (partie iv) ■ Graphiques - Exactitude - Tests sur 2 proportions - Tailles n_1 et n_2 inégales (seuil $\alpha = 0,05$)

(Ex14)

(Ex15)

(Ex16)

(Ex17)

(Ex18)

Figures Ex (partie v) ■ Tests sur 2 proportions - Tailles inégales (seuil $\alpha = 0,01$)

(Ex19)

(Ex20)

(Ex21)

(Ex22)

(Ex23)

Figures Ex (partie vi) ■ Tests sur 2 proportions - Tailles inégales (seuil $\alpha = 0,01$)

(Ex24)

(Ex25)

(Ex26)

(Ex27)

(Ex28)

Figures Ex (partie vii) ■ Tests sur 2 proportions - Tailles inégales (seuil $\alpha = 0,01$)

(Ex29)

(Ex30)

(Ex31)

(Ex32)

(Ex33)

Figures Ex (partie viii) ■ Tests sur 2 proportions - Tailles inégales (seuil $\alpha = 0,01$)

(Ex34)

(Ex35)

(Ex36)

(Ex37)

(Ex38)

Figures Ex (partie ix) ■ Tests sur 2 proportions. Les données des graphiques de droite sur tailles inégales sont basées sur des tests de différence à variance d'erreur basée sur la moyenne harmonique des n (≈ 17), la figure EX38 pour tailles égales épousant cette valeur.

(Ex39)

(Ex40)

(Ex41)

(Ex42)

(Ex43)

Figures Pu (partie i) ■ Puissance. Tests sur 2 proportions - Tailles égales ($\alpha = 0,05$). Les 27 graphiques suivants présentent les courbes de puissance de nos cinq tests de la différence entre deux proportions (en mode unilatéral, à droite), ce pour des échantillons de tailles variables et des conditions paramétriques différentes. Les deux échantillons comparés sont formés à partir des binomiales $x_1 \sim B(n_1, \pi_1)$ et $x_2 \sim B(n_2, \pi_2)$, où π_1 est la valeur inscrite en abscisse et π_2 la valeur réduite telle qu'indiquée dans le titre du graphique. Plusieurs graphiques présentent des lignes superposées, réunissant des fonctions « y » sous la couleur de la dernière tracée (souvent le « jaune » de la fonction $y_{F-T/2}$).

(Pu1)

(Pu2)

(Pu3)

(Pu4)

(Pu5)

(Pu6)

Figures Pu (partie ii) ■ Puissance. Tests sur 2 proportions - Tailles égales ($\alpha = 0,05$) (suite).

(Pu7)

(Pu8)

(Pu9)

(Pu10)

(Pu11)

Figures Pu (partie iii) ■ Puissance. Tests sur 2 proportions - Tailles inégales ($\alpha = 0,05$).

(Pu10)

(Pu11)

(Pu12)

(Pu13)

(Pu14)

(Pu15)

Figures Pu (partie iv) ■ Tests sur 2 proportions - Tailles inégales ($\alpha = 0,05$) (suite).

(Pu16)

(Pu17)

(Pu18)

(Pu19)

(Pu20)

(Pu21)

Figures Pu (partie v) ■ Tests sur 2 proportions - Tailles inégales ($\alpha = 0,05$) (suite).

(Pu22)

(Pu23)

(Pu24)

(Pu25)

(Pu26)

(Pu27)