7. Comparaison des moyens traitements avec un contrôle :

Dans beaucoup d'expériences, un des traitements est considéré comme un contrôle et l'analyse d'intérêt est de comparer chacun des $a-1$ traitements moyens (restant) avec le traitement de contrôle. On a, ainsi, $a-1$ comparaisons à faire seulement. On définit la procédure de comparaison de Dunnett (1964) où on suppose que le traitement a est le contrôle et on teste le corps d'hypothèses suivant :

$$\begin{cases} H_0 : \mu_i = \mu_a, \quad \forall\, i = 1, 2, \cdots, a-1 \\ H_a : \mu_i \neq \mu_a \end{cases}$$

La procédure de Dunnett-Test est une approche de modification du $t - test$ usuel. Pour chaque hypothèse, on estime/calcule les différences observée de l'échantillon des moyennes, i.e.

$$|\bar{y}_i - \bar{y}_a| \quad \forall\, i = 1, 2, \cdots, a-1$$

$$Rejet\ de\ l'hypothèse\ de\ base, H_0 : \mu_i = \mu_a, pour\ un\ seuil\ de\ signification\ \alpha\ si$$

$$|\bar{y}_i - \bar{y}_a| > d_\alpha(a-1, f)\sqrt{CM_{SS_E}\left(\frac{1}{n_i} + \frac{1}{n_a}\right)}$$

$$où\ la\ constante - valeur\ critique, d_\alpha(a-1, f), est\ lue\ sur\ la\ table\ du\ Dunnett'test;$$
$$Notons\ que\ tests\ bilatéral\ et\ unilatéral\ sont\ possibles\ pour\ Dunnett - procédure.$$

*Remarque : Au niveau de la comparaison des traitements avec un contrôle, il serai bien utile d'utiliser une taille élevée, n_a, pour le traitement de contrôle que les $a-1$ autres traitements à taille égales, n . On choisit approximativement le ratio $\frac{n_a}{n} = \sqrt{a}$.

8. Détermination de la taille d'échantillon :

Un des problèmes majeurs du plan d'expérience est la décision critique du choix de la taille d'échantillon, i.e. de la détermination du nombre de répétitions/ répliques à exécuter[1]. Plusieurs approches sont définies :

i. Courbes caractéristiques « fonctionnelles» : qu'on définit comme plot de la probabilité de l'erreur de type II d'un test statistique défini pour une taille particulière d'échantillon contre un paramètre reflétant la mesure pour laquelle l'hypothèse de base est rejetée.

Ces courbes sont utiles pour guider l'expérimentateur à sélectionner le nombre de répliques au niveau duquel se définit la sensibilité du plan aux différences importantes dans les traitements.

[1] Plus de répliques nécessaires pour l'expérimentateur intéressé par la détection d'effets importants des traitements.

On considère la probabilité de l'erreur de type II d'un modèle statistique à effets fixes, pour le cas où les traitements sont à tailles-échantillons égales comme suit,

$$\beta = 1 - P[Rejet\ de\ H_0/H_0\ fausse]$$
$$= 1 - P\left[F_0 > F_{\alpha,a-1,N-a}/H_0\ fausse\right]$$
$$O\grave{u}\ l'\acute{e}valuation\ de\ la\ probabilit\acute{e}\ n\acute{e}cessite\ la\ connaissance\ de\ la\ distribution$$
$$de\ la\ statistique\ du\ test\ F_0\ avec\ \ H_0\ fausse.$$
$$\Longrightarrow$$
$$Si\ \ H_0\ est\ fausse\ alors\ F_0 = \frac{CM_{SS_{traitements}}}{CM_{SS_E}} \backsim \mathcal{F}(a-1, N-a, \lambda)$$
$$o\grave{u}\ \ \lambda = \frac{\mathrm{n}\sum_{\mathrm{i}=1}^{\mathrm{a}} \alpha_{\mathrm{i}}^2}{\sigma^2}\ \ param\grave{e}tre\ de\ non\ centralit\acute{e}.$$
$$Pour\ \lambda = 0, on\ retrouve\ la\ distribution\ de\ Fisher, F, usuelle\ (centrale).$$

Et courbes caractéristiques fonctionnelles définissent le plot de la probabilité de l'erreur de type II (β) (i. e. le pouvoir du test $(1 - \beta)$) contre un paramètre ϕ, où

$$\phi^2 = \frac{\mathrm{n}\sum_{\mathrm{i}=1}^{\mathrm{a}} \alpha_{\mathrm{i}}^2}{\mathrm{a}\sigma^2}\ \text{, grandeur liée au paramètre de non centralité } \lambda;$$

Courbes caractéristiques définies pour seuil $\alpha = 5\%$ (ou 1%) et un rang de ddl v_1 et v_2 de Fisher.

ii. <u>Spécification d'une augmentation de la déviation standard</u>, σ : Méthode occasionnellement aidant à choisir la taille de l'échantillon n.

Dans le cas où les moyennes traitements ne diffèrent pas, alors la déviation standard d'une observation choisie au hasard est définie par σ.

Dans le cas contraire où les moyennes traitements diffèrent, la déviation standard d'une observation choisie au hasard est définie par

$$\sqrt{\sigma^2 + \left(\sum_{\mathrm{i}=1}^{\mathrm{a}} \alpha_{\mathrm{i}}^2 / \mathrm{a}\right)}$$

Pour un pourcentage P d'augmentation de l'écart type d'une observation au-delà duquel on rejette l'hypothèse de base où toutes les moyennes traitements sont égales, on définit

$$\frac{\sqrt{\sigma^2 + \left(\sum_{i=1}^{a} \alpha_i^2 / a\right)}}{\sigma} = 1 + 0.01P$$

or,

$$\frac{\sqrt{\sum_{i=1}^{a} \alpha_i^2 / a}}{\sigma} = \sqrt{(1 + 0.01P)^2 - 1}$$

$$\Rightarrow \ \phi = \frac{\sqrt{\sum_{i=1}^{a} \alpha_i^2 / a}}{\sigma / \sqrt{n}} = \sqrt{(1 + 0.01P)^2 - 1} \ \left(\sqrt{n}\right)$$

Pour une valeur spécifiée de P, on estime le paramètre ϕ et on utilise les Courbes caractéristiques « fonctionnelles» pour déterminer la taille n exigée de l'échantillon.

iii. <u>Méthode d'estimation par Intervalle de Confiance :</u>

Approche qui suppose que l'expérimentateur souhaite exprimer les résultats finaux en termes d'intervalles de confiance et spécifier à l'avance quelle largeur il souhaite que ces intervalles de confiance soient.

On définit la précision de l'intervalle par

$$\pm t_{\alpha/2, N-a} \sqrt{\frac{2CM_{SS_E}}{n}}$$

n définit la taille d'échantillon faible qui mène à la précision désirée.

Exemple : Soit $\sigma^2 = 3^2$une estimation de CM_{SS_E}.

On essaye avec trois valeurs de repliques n pour trouver la précision désirée de l'IC(.):

$$Pour\ n = 5\ répliques \ \Rightarrow \ \pm t_{\alpha/2, N-a} \sqrt{\frac{2CM_{SS_E}}{n}} = \pm 2.086 \sqrt{\frac{2(9)}{5}} = \pm 3.96$$

$$Pour\ n = 4\ répliques \ \Rightarrow \ \pm t_{\alpha/2, N-a} \sqrt{\frac{2CM_{SS_E}}{n}} = \pm 2.136 \sqrt{\frac{2(9)}{4}} = \pm 4.52$$

$$Pour\ n = 3\ répliques \ \Rightarrow \ \pm t_{\alpha/2, N-a} \sqrt{\frac{2CM_{SS_E}}{n}} = \pm 2.228 \sqrt{\frac{2(9)}{3}} = \pm 5.46$$

Soit n = 4 définit la taille minimale de l'échantillon qui détermine la décision désirée.

9. Méthode non paramétrique pour l'analyse de variance :

i. La procédure de Kruskal-Wallis Test : Dans les situations où l'hypothèse de normalité n'est pas vérifiée, l'expérimentateur essaye d'utiliser une procédure alternative à l'analyse $F-test$ de la variance qui ne dépend pas de la normalité. On définit, ainsi, une procédure de test développée par Kruskal-Wallis (1952) pour tester l'hypothèse de base que les a traitements sont identiques vs l'hypothèse alternative que certains traitements génèrent des observations plus larges que d'autres. La méthode de Kruskal-Wallis-Test est une alternative[2] non paramétrique à l'analyse usuelle de la variance.

Pour l'approche de Kruskal-Wallis-Test, on classe d'abord les observations par ordre croissant et on remplace chaque observation par son rang, R_{ij}, de la plus petite observation ayant le rang 1 ; En cas d'égalité (i.e. les observations ayant la même valeur), attribuer le rang moyen à chacune des observations égales, ex aequo.

Soit $R_{i.}$ la somme des rangs dans le ième traitement, la statistique du test est définie par

$$H = \frac{1}{S^2}\left[\sum_{i=1}^{a} \frac{{R_{i.}}^2}{n_i} - \frac{N(N+1)^2}{4}\right] \qquad (1)$$

où

n_i nombre d'observations dans le ième traitement et N nombre total d'observations.

$$\text{La variance des rangs } S^2 = \frac{1}{N-1}\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4}\right]$$

$$** \text{ S'il n'ya pas de liens d'égalité, } S^2 = \frac{N(N+1)}{12}$$

Et la statistique du test se simplifie à l'écriture suivante:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{a} \frac{{R_{i.}}^2}{n_i} - 3(N+1) \qquad (2)$$

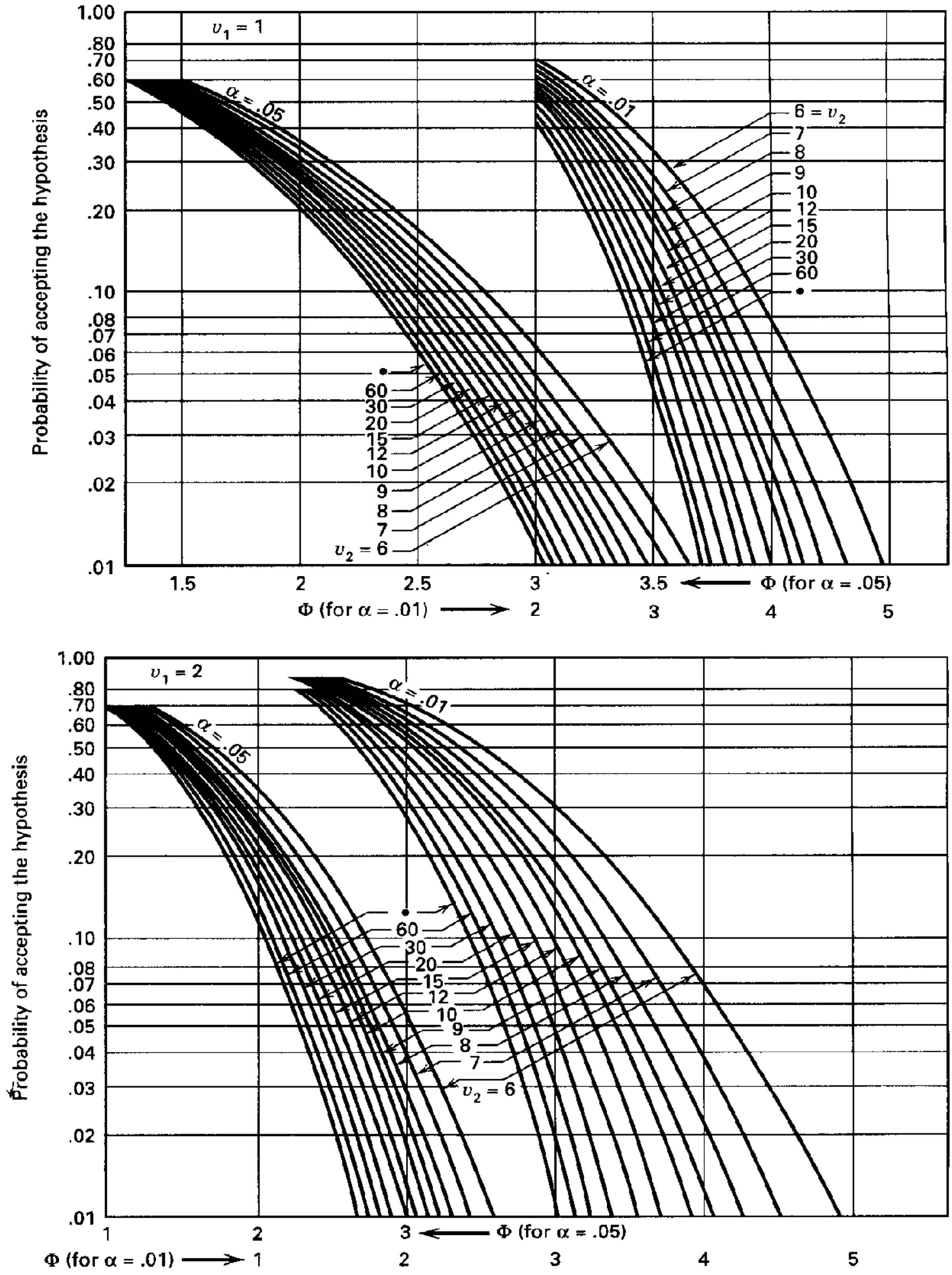
lorsque le nombre de liens est modéré (diminué), il y aura une faible différence entre équations (1) et (2) et la forme simplifiée peut être utilisée.

Dans le cas où les n_i sont raisonnablement larges , soit $n_i \geq 5$, alors sous H_0 , la statistique $H \xrightarrow[\text{asymthotiquement}]{} \chi^2_{a-1}$

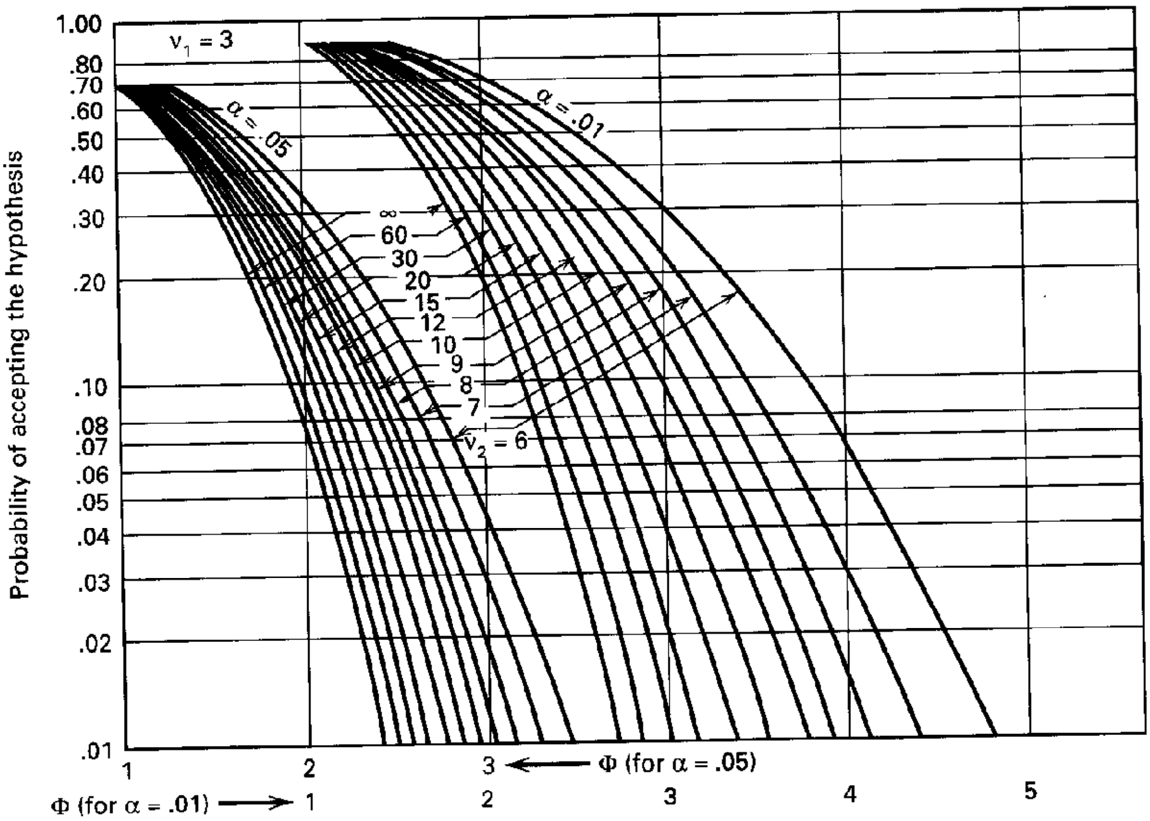
et on a rejet de H_0 pour $H > \chi^2_{\alpha,a-1}$ et Pvalue $< \alpha\%$

[2] En effet, étant donné que la procédure est conçue être sensible pour tester les différences de moyennes, il est parfois pratique de considérer la procédure de Kruskal-Wallis-test comme un test d'égalité des moyennes de traitement.

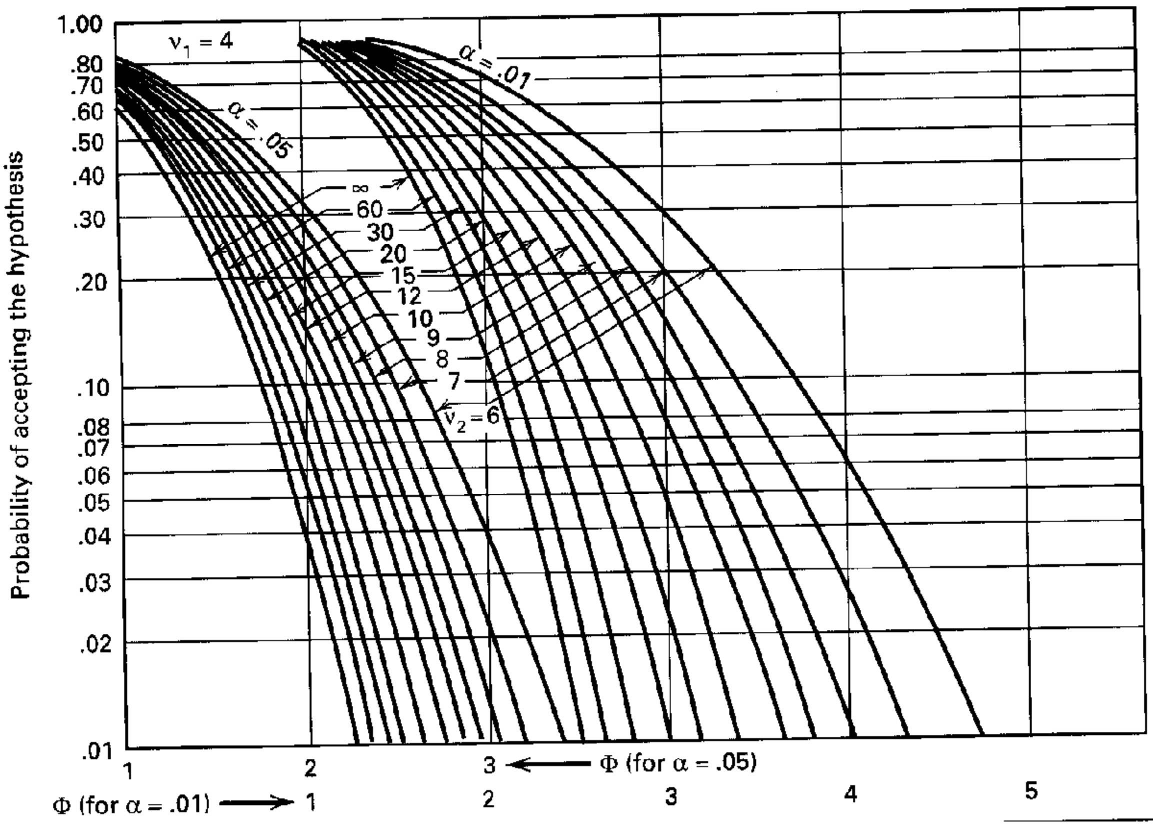
Remarque : Aperçu sur les Courbes caractéristiques « fonctionnelles» d'un modèle à effets fixes de l'analyse de variance :



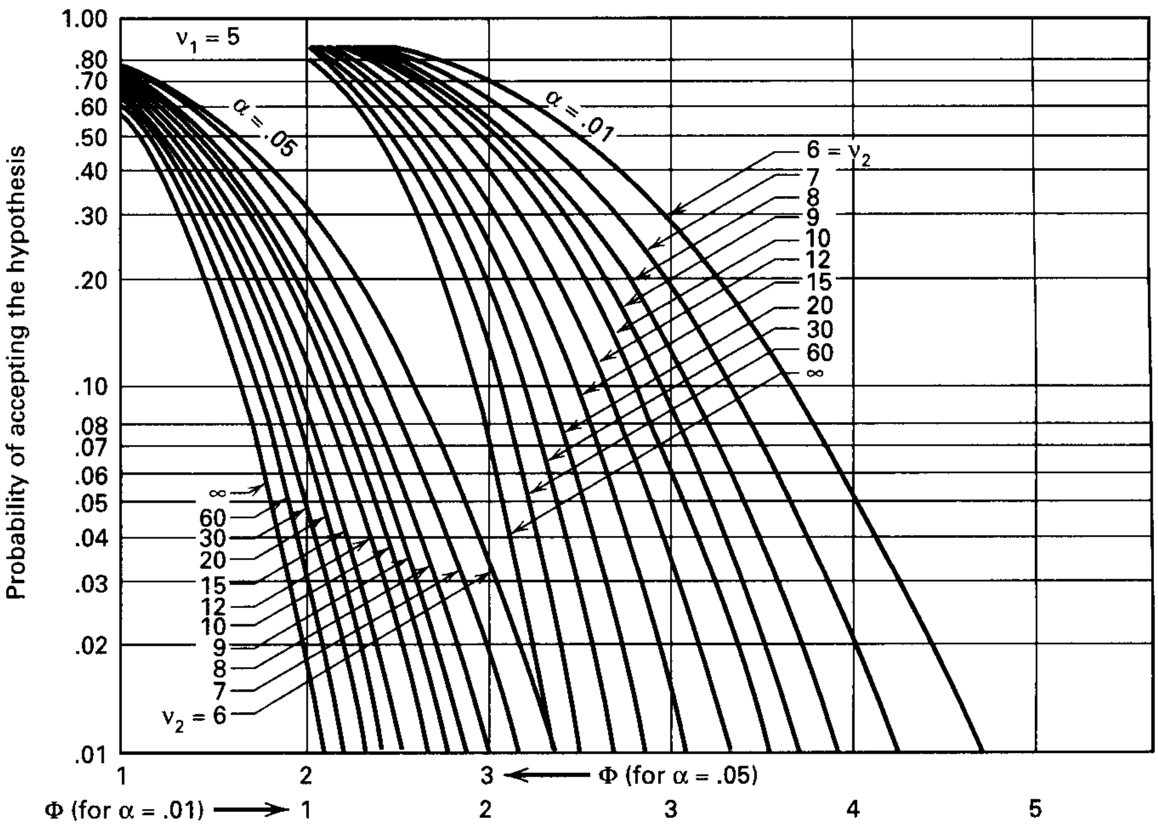
v_1 = numerator degrees of freedom, v_2 = denominator degrees of freedom

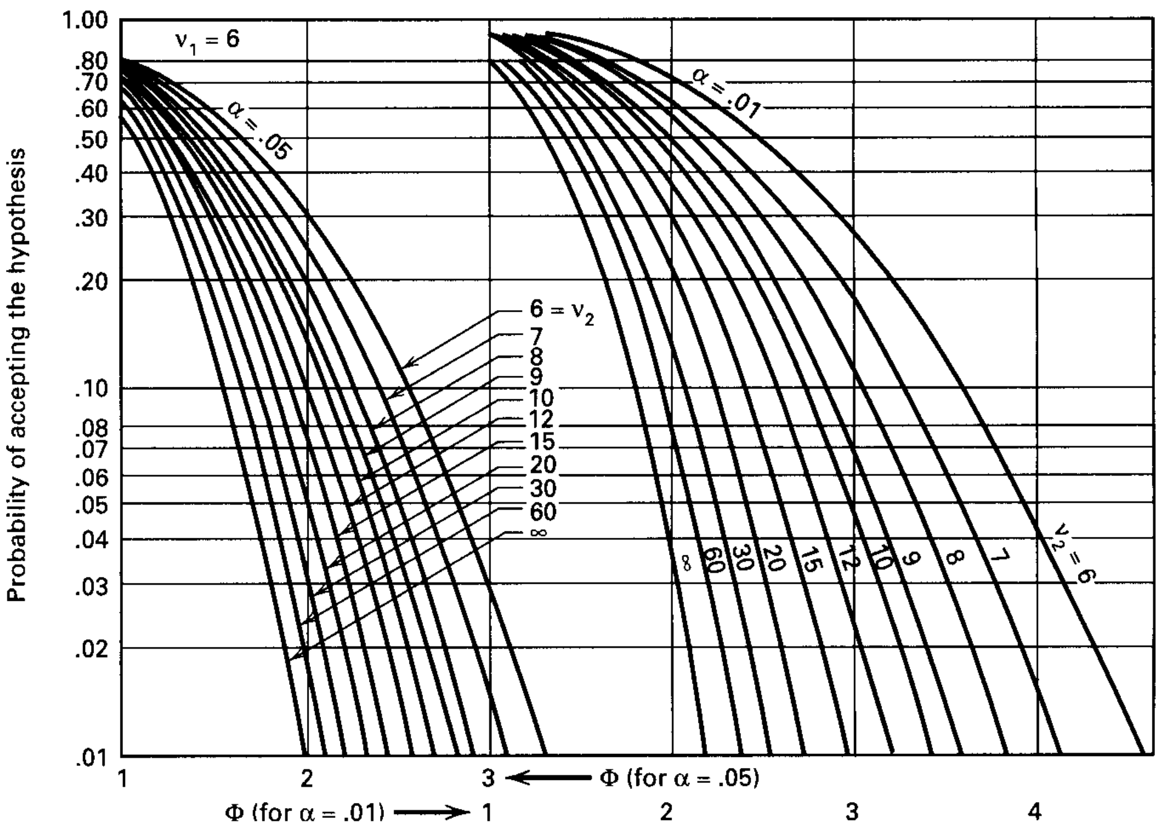


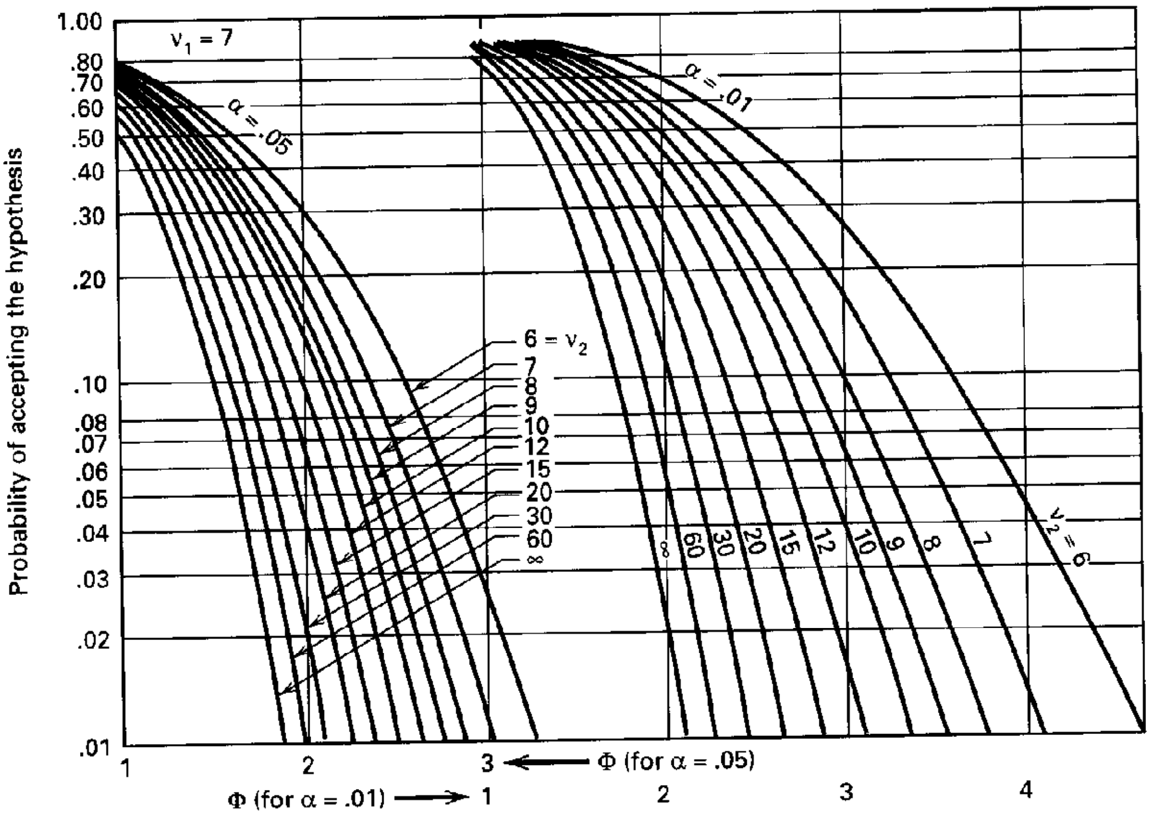
$\nu_1 = 3$
$\alpha = .05$
$\alpha = .01$
Probability of accepting the hypothesis
1.00
.80
.70
.60
.50
.40
.30
.20
.10
.08
.07
.06
.05
.04
.03
.02
.01
∞
60
30
20
15
12
10
9
8
7
$\nu_2 = 6$
1
2
3
Φ (for α = .05)
Φ (for α = .01)
1
2
3
4
5



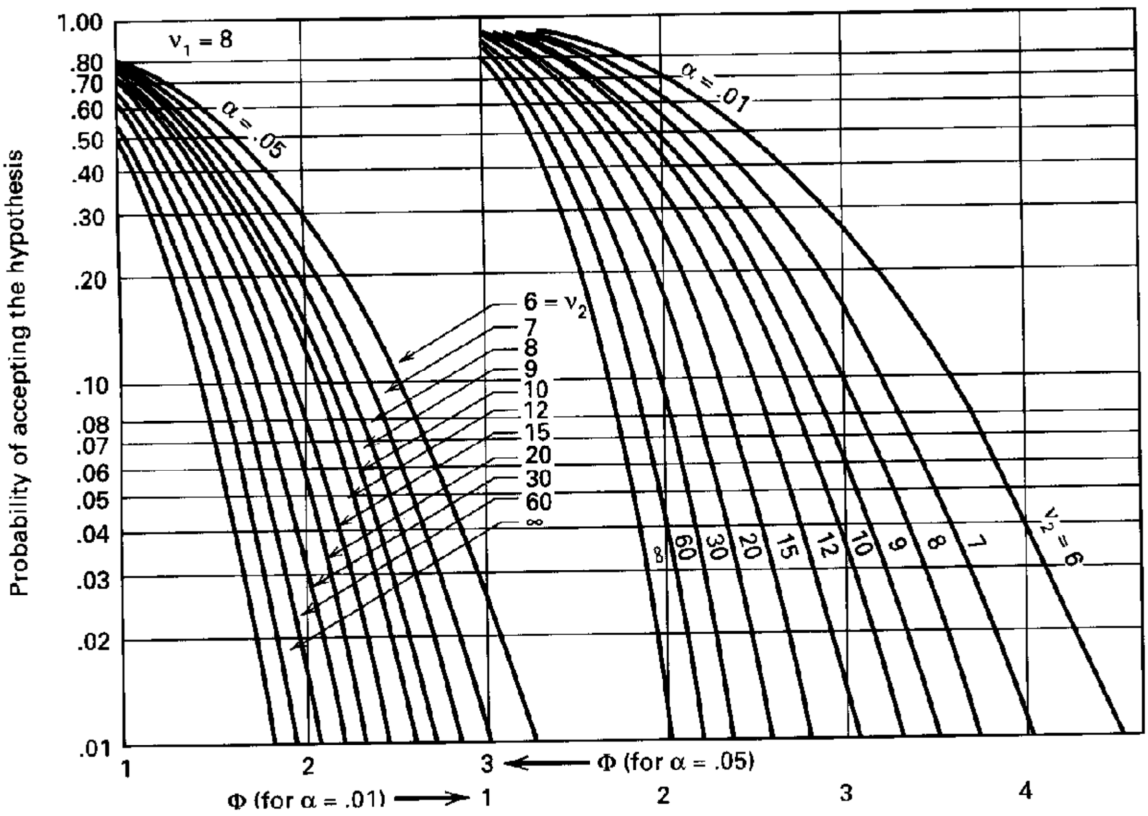
$\nu_1 = 4$
$\alpha = .05$
$\alpha = .01$
Probability of accepting the hypothesis
1.00
.80
.70
.60
.50
.40
.30
.20
.10
.08
.07
.06
.05
.04
.03
.02
.01
∞
60
30
20
15
12
10
9
8
7
$\nu_2 = 6$
1
2
3
Φ (for α = .05)
Φ (for α = .01)
1
2
3
4
5


$\nu_1 = 5$
$\alpha = .05$
$\alpha = .01$
Probability of accepting the hypothesis
1.00
.80
.70
.60
.50
.40
.30
.20
.10
.08
.07
.06
.05
.04
.03
.02
.01
$6 = \nu_2$
7
8
9
10
12
15
20
30
60
∞
∞
60
30
20
15
12
10
9
8
7
$\nu_2 = 6$
1
2
3
Φ (for α = .05)
Φ (for α = .01)
1
2
3
4
5


$\nu_1 = 6$
$\alpha = .05$
$\alpha = .01$
Probability of accepting the hypothesis
1.00
.80
.70
.60
.50
.40
.30
.20
.10
.08
.07
.06
.05
.04
.03
.02
.01
$6 = \nu_2$
7
8
9
10
12
15
20
30
60
∞
∞
60
30
20
15
12
10
9
8
7
$\nu_2 = 6$
1
2
3
Φ (for α = .05)
Φ (for α = .01)
1
2
3
4



$\nu_1 = 7$
Probability of accepting the hypothesis
1.00
.80
.70
.60
.50
.40
.30
.20
.10
.08
.07
.06
.05
.04
.03
.02
.01
$\alpha = .05$
$\alpha = .01$
$6 = \nu_2$
7
8
9
10
12
15
20
30
60
∞
$\nu_2 = 6$
1
2
3
4
Φ (for $\alpha = .05$)
Φ (for $\alpha = .01$)



$\nu_1 = 8$
Probability of accepting the hypothesis
1.00
.80
.70
.60
.50
.40
.30
.20
.10
.08
.07
.06
.05
.04
.03
.02
.01
$\alpha = .05$
$\alpha = .01$
$6 = \nu_2$
7
8
9
10
12
15
20
30
60
∞
$\nu_2 = 6$
1
2
3
4
Φ (for $\alpha = .05$)
Φ (for $\alpha = .01$)