RÉPONSE DE SCGM À UNE DEMANDE DE RENSEIGNEMENT

Origine : Demande de renseignement en date du 28 juin 2001

Demandeur : CERQ

Référence : SCGM-1, document 03, page 22 de 27, Le facteur de déplacement, graphique

Questions :

5.1 Comme le graphique de la page 22 de 27 repose sur l'hypothèse que la distribution des variations de volume en gaz de réseau est « approximativement normale », est-ce que cette hypothèse correspond de près à la réalité observable (donner les détails et les résultats d'un test à faire à ce sujet) ?

5.2 Si la réponse apportée à la question 5.1 indique l'existence d'un écart significatif par rapport à « une distribution approximativement normale », quel est l'impact de cet écart sur les résultats du tableau 2 et du graphique de la page 11 de 27, SCGM-1, dot. 3, de même que sur le graphique de la page 22 de 27, SCGM-1, doc. 3 ? Il serait important de détailler votre réponse.

5.3 Si la réponse apportée à la question 5.1 indique l'existence d'un écart significatif par rapport à « une distribution approximativement normale », quel serait l'impact global d'une telle situation sur le nouveau programme de produits financiers dérivés proposé par SCGM ?

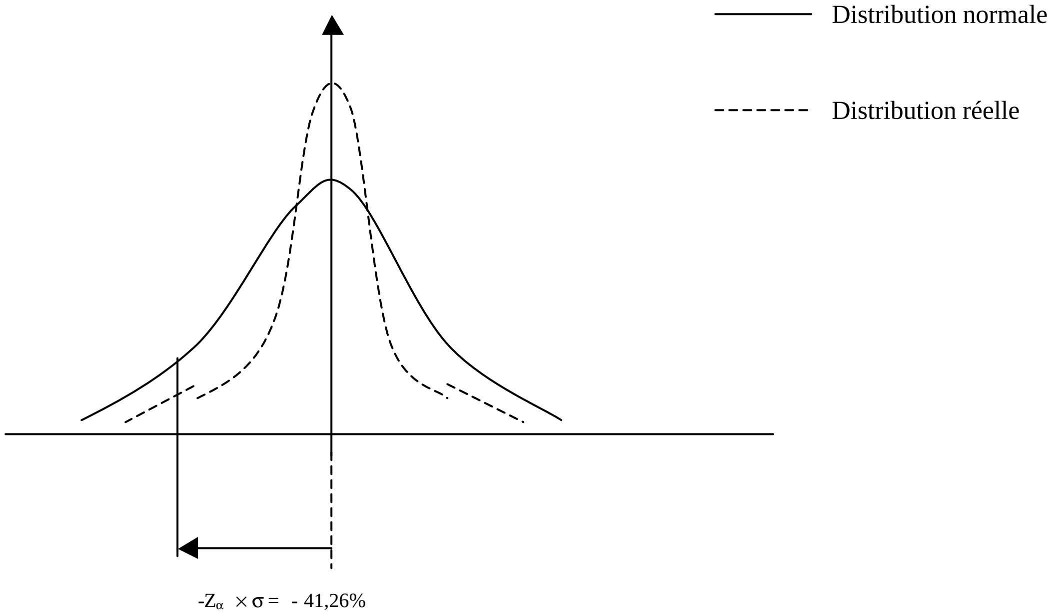
Réponse :

5.1

Il est difficile de prévoir la valeur précise des variations annuelles du gaz de réseau. SCGM a donc proposé des hypothèses qu'elle croit conservatrices, notamment quant aux caractéristiques principales de la distribution de probabilité. D'abord, il est clair que les variations peuvent être aussi bien positives que négatives. Comme il n'y a présentement pas de raison de s'attendre à une croissance plutôt qu'à une décroissance, la première hypothèse concernant la variation du gaz de réseau est que son espérance mathématique égale 0. Par ailleurs, SCGM s'attend à trouver des variations faibles beaucoup plus souvent que des variations élevées, de façon à ce que la distribution soit unimodale. En vertu de ce qui précède, SCGM a présumé que les variations suivent une distribution normale d'espérance nulle ($\mu = 0$). Compte tenu de l'importance conférée à cette hypothèse, il convient de procéder à des tests de normalité sur la variation des volumes de gaz de réseau.

1. Forme de l'histogramme:

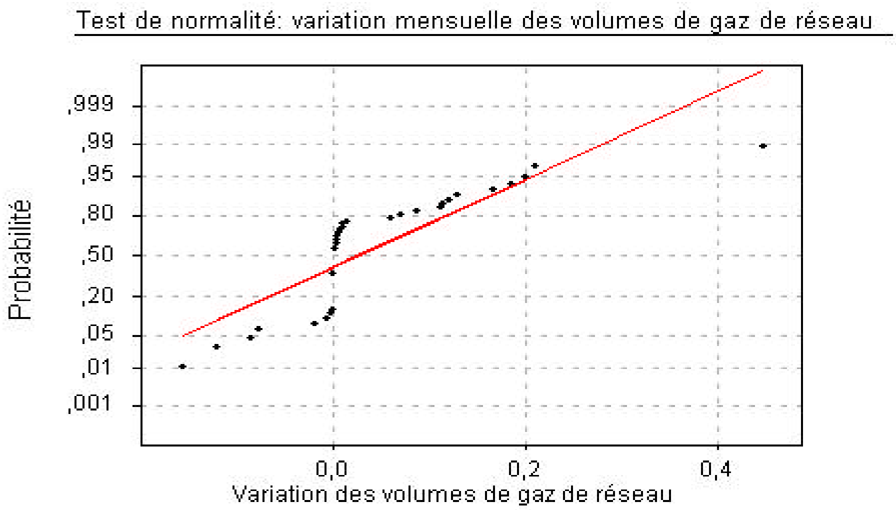
   De prime abord, il semble tout à fait naturel d'examiner visuellement la parité entre la distribution théorique, en l'occurrence la loi normale, et la distribution observée de l'échantillon étudié. La forme de l'histogramme révèle une distribution aux courbes plus étroites que l'allure classique en forme de cloche de la loi normale.



2. Ajustement graphique :

   En second lieu, il s'agit d'étudier la fonction de répartition de l'échantillon. En effet, pour la plupart des lois de probabilité, et en particulier pour la loi normale, une transformation fonctionnelle simple permet de représenter la fonction de répartition de la distribution par une droite[1.] Il s'agit donc de vérifier l'adéquation de la fonction de répartition empirique de l'échantillon à une droite (fonction de répartition de la loi normale) sur un graphique à échelle fonctionnelle. Le graphique ci-dessous indique que la fonction de répartition basée sur les données observées s'éloigne de la droite représentant la loi normale théorique.

1 La fonction de répartition correspond `ala fonction de distribution cumulative.



3. Test de Kolmogorov-Smirnoff :

La dernière méthode proposée est le test de Kolmogorov-Smirnoff. Il s'agit d'un test d'hypothèse non-paramétrique visant à déterminer si l'éloignement entre la distribution inconnue et la distribution théorique est significatif ou non. L'élaboration de ce test statistique comprend les étapes suivantes :

1ère étape : Énoncer les hypothèses de base.

2^{ième} étape : Calculer la statistique de Kolmogorov-Smirnoff (K-S)

3^{ième} étape : Déterminer la région critique : établir la valeur théorique de la statistique tirée de la table de K-S.

4^{ième} étape : Comparer les deux valeurs et conclure.

1ère étape : Hypothèses statistiques :

La première étape de la méthodologie consiste à établir les hypothèses statistiques. Dans le cadre de cette étude, les hypothèses nulles et alternatives, désignées respectivement par H_0 et H_1 s'énoncent comme suit :

H_0: La variable aléatoire est normalement distribuée.

H_1 : La variable aléatoire ne suit pas une loi normale.

En d'autres termes :

$H_0 : F(x) = F_0(x)$

$H_1 : F(x) \neq F_0(x)$

Où :

F(x) est la fonction de répartition de la variable échantillonnée

$F_0(x)$ est la fonction de répartition hypothétique (de loi normale)

2^{ième} étape : Calcul de la statistique de Kolmogorov-Smirnoff

Il faut ensuite spécifier la statistique sur laquelle portera le test. Dans les processus non paramétriques, l'objectif de cette statistique est généralement de quantifier l'éloignement de la distribution d'échantillonnage par rapport à la distribution théorique. Pour ce faire, il suffit de scinder les fonctions de densité en sous-intervalles. Cette partition permet d'obtenir les distributions de fréquences observées et espérées et, par sommation, les distributions de fréquences cumulatives espérées et observées. Ces détails sont présentés dans le tableau suivant :

Tableau des fréquences

Classes	1	2	3	….	C
Fréquence espérée	E_1	E_2	E_3	....	E_C
Fréquence observée	O_1	O_2	O_3	....	O_C
Fréquence cumulative espérée	F_{E1}	F_{E2}	F_{E3}	....	F_{EC}
Fréquence cumulative observée	F_{O1}	F_{O2}	F_{O3}	....	F_{OC}

Où :

C : Nombre de classes.

O_C : Nombre d'observations dans la classe C.

E_C : Nombre d'observations espérées.

$E_C = n \times p_c$

n : Nombre total d'observations dans l'échantillon.

p_c : Probabilité qu'une observation de l'échantillon se trouve dans la classe c lorsque l'hypothèse H_0 est vraie.

La statistique de Kolmogorov mesure l'écart maximal entre la fréquence cumulative observée et la fréquence cumulative espérée, normalisée par le nombre d'observations. De façon succincte, la statistique de Kolmogorov se formule comme suit :

$$D = \frac{MAX\left|F_0 - F_E\right|}{n}$$

Où :

D est la variable de Kolmogorov-Smirnoff

F_E correspond aux fréquences cumulatives tirées de la distribution hypothétique (distribution espérée si l'hypothèse H_0 était vraie)

F_O correspond aux fréquences cumulatives tirées de la distribution observée (de l'échantillon).

n représente la taille de notre échantillon.

$3^{\text{ième}}$ étape: La règle de décision

L'étape finale consiste à définir le processus qui mène à la conclusion. Cette conclusion se base sur l'estimé de la statistique de Kolmogorov et sur la valeur critique tirée de la table de Kolmogorov. La règle de décision est la suivante :

accepter l'hypothèse nulle (non rejet de H_0), si $D < D_c$:

rejeter l'hypothèse nulle autrement

Il est à noter que la confirmation ou l'infirmation de l'hypothèse nulle est toujours faite avec une certaine probabilité et qu'en ce sens, le test ne donne aucune certitude absolue. Autrement dit, l'acceptation de l'hypothèse nulle ne signifie pas pour autant que la variable aléatoire étudiée suit une loi normale, mais plutôt que cette dernière est une approximation relativement raisonnable de la distribution.

Résultats empiriques :

Cette section énonce les résultats de la méthodologie de Kolmogorov-Smirnoff présentée ci-dessus.

Il convient d'abord de noter les points suivants :

- La méthode a été appliquée sur un échantillon de 42 observations.
- Le paramètre a associé à la zone de rejet est fixé à 5%.

Le cadre analytique étant précisé, il est maintenant possible d'énumérer les résultats:

La valeur critique (D_c) de Kolmogorov associée à la zone de rejet est calculée comme suit :

$$D_{0,05} = \frac{1{,}36}{\sqrt{n}} \quad \text{où } n=42,$$

$D_{\mathbf{c}} = 0{,}210$

Le test sur l'échantillon indique :

$D = 0{,}321$

Donc :

$D_c < D$

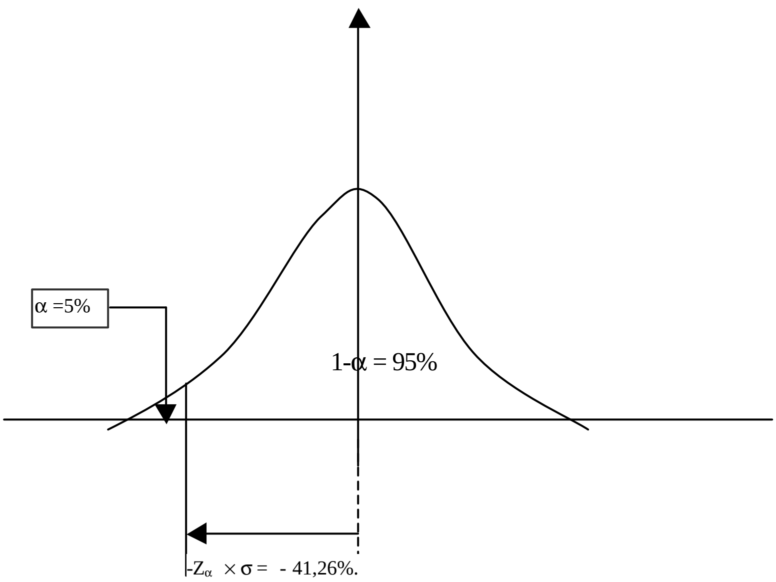
Comme la valeur de l'échantillon est supérieure à la valeur critique, il faut conclure que les variations mensuelles de volume de gaz de réseau ne suivent pas une loi normale confirmant ainsi les tests précédents.

5.2

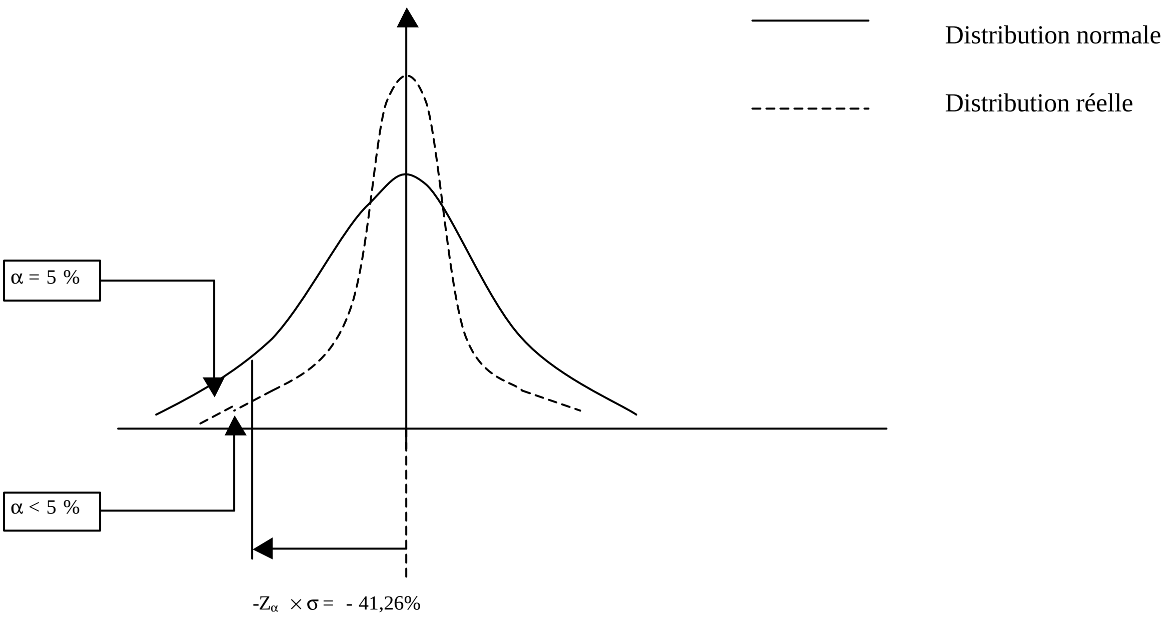
Comme déjà mentionné, SCGM s'est montrée conservatrice dans l'établissement de ces hypothèses, notamment celle concernant la normalité des variations. De fait, il appert que la conséquence de notre approximation affecte à la baisse la valeur du taux de substitution des tableaux de la page 11 et 22. Voici pourquoi:

Il faut d'abord rappeler que le taux de substitution, établi à 41,26%, résulte du produit de l'écart type annuel de la variation des volumes de gaz de réseau (σ_{annuel} = 25,01%) et de la variable aléatoire normale centrée réduite Z_α associée au niveau de confiance choisi à 95% (Z =1,65). Dans la mesure où la distribution des variations des volumes de gaz de réseau obéit à une loi normale, il existe une probabilité de 1-α = 95% que le taux de substitution soit supérieur à la borne $-Z_\alpha \times \sigma$ = - 41,26% (ou inférieur à 41,26%). Cette probabilité correspond à l'aire sous la courbe de la distribution à droite de la borne de 41% tel que présenté sur le graphique suivant :

<u>Distribution approximativement normale des variations de volume de gaz</u>



La superposition de la distribution réelle à la distribution théorique normale permet de constater que la borne établie à un niveau de confiance de 1- α = 95 % se trouve, dans le cas de cette dernière, beaucoup plus en queue de distribution. Conséquemment, l'aire sous la courbe de la distribution réelle à gauche de la borne (α < 5 %) est bien inférieure à l'aire correspondante dans le cas de la distribution normale (α = 5 %). Autrement dit, la probabilité que le taux de substitution soit supérieur à la borne $-Z_\alpha \times \sigma$ = - 41,26% dans le cas de la distribution réelle est dans les faits supérieure à 95%.



L'hypothèse de normalité, même si elle correspond à une approximation, permet donc d'adopter une approche plus prudente.

5.3

Cela n'affecte en rien le nouveau programme de produits dérivés. En réalité, cela devrait réduire davantage les limites volumétriques annuelles puisque le taux de substitution utilisé est supérieur au taux de substitution réel pour le niveau de confiance choisi.