VAR à densité

Le théorème de transfert est une appellation qui recouvre de nombreux énoncés ayant un point commun : transférer un changement de variable “aléatoire" dans un calcul d'intégrale ou de somme.
Ces différents énoncés sont les manifestations d'une même idée selon les contextes. Sans justifier de ce qui les associe, nous admettrons que le présent supplément regroupe bien sous une même dénomination des résultats très similaires mais formulés selon le contexte d'utilisation des variables aléatoires à densité.

Théorème de transfert : cas continu

> **Théorème de transfert (*formulation ex-ECS*)** : Soit X une variable aléatoire ayant une densité f_X nulle en dehors de l'intervalle $]a;b[$, avec $-\infty \leq a < b \leq +\infty$ et soit g une fonction continue sur $]a;b[$ éventuellement privé d'un nombre fini de points.
> $\mathbb{E}[g(X)]$ existe et est égale à $\displaystyle\int_a^b g(t) f_X(t)\ \mathrm{d}t$ si et seulement si cette intégrale converge absolument.

Exemple 1 : On donne X dont la densité est $f_X(t) = 4t^3$ sit $t \in [0;1[$ et nulle en dehors. On souhaite déterminer $\mathbb{E}[\sqrt{X}]$.
On a alors, dans ce cas, $a = 0$ et $b = 1$ et on travaille sur $I =]0;1[$. La fonction $g : t \mapsto \sqrt{t}$ est bien définie sur $]0;1[$ et on peut calculer $\displaystyle\int_0^1 \sqrt{t} \cdot 4t^3\ \mathrm{d}t$ comme une intégrale sur un segment d'une fonction continue sur $[0;1]$ (par produit), ce qui assure la convergence de l'intégrale (faussement impropre) et on calcule ainsi :

$$\mathbb{E}[g(X)] = \mathbb{E}[\sqrt{X}] = \int_0^1 4t^{\frac{7}{2}}\ \mathrm{d}t = \left[\frac{2\times 4}{9} t^{\frac{9}{2}}\right]_0^1 = \frac{8}{9}$$

Démonstration : L'ancien programme ECS propose de limiter la démonstration au cas où g est de classe $\mathcal{C}^1$ sur $]a;b[$ vérifiant g' strictement positive (ou strictement négative). De plus, cette démonstration, même restreinte à un cas particulier, n'est pas exigible dans ce programme.

Si l'on examine la variable aléatoire $Y = g(X)$ et que l'on suppose qu'elle admet une espérance, alors on peut écrire, par définition :

$$\mathbb{E}[Y] = \int_a^b x f_Y(x)\ \mathrm{d}x$$

les bornes a et b étant bien obtenues comme bornes de $g(I)$ où $X : \Omega \longrightarrow \mathbb{R}$, en notant $I = X(\Omega)$ (en effet, g est continue donc l'image de I est un intervalle par le théorème des valeurs intermédiaires). On va donc procéder à un changement de variables $x = g(t)$ permettant d'écrire $\mathrm{d}x = \mathrm{d}g(t) = g'(t)\ \mathrm{d}t$ et $x = a \iff t = g^{-1}(a)$ et $x = b \iff t = g^{-1}(b)$ pour changer les bornes, ayant g' strictement positive on aura g strictement croissante donc bijective de $I =]\alpha;\beta[$ dans l'intervalle $]a;b[= g(I)$ avec $g(\alpha) = a$ et $g(\beta) = b$. On réécrit alors :

$$\mathbb{E}[Y] = \mathbb{E}[g(X)] = \int_\alpha^\beta g(t) g'(t) f_Y(g(t))\ \mathrm{d}t$$

Il resterait donc à établir le lien entre la densité f_X de X et celle f_Y de Y. Pour ce faire, passons par les fonctions de répartition :

$$\forall x \in]a;b[\ \int_{-\infty}^x f_Y(t) = F_Y(x) = \mathbb{P}[Y \leq x] = \mathbb{P}[g(X) \leq x] = \mathbb{P}[X \leq g^{-1}(x)] = \int_{-\infty}^{g^{-1}(x)} f_X(t)\ \mathrm{d}t$$

et ainsi, en rappelant que les fonctions de répartition sont de limite nulle en $-\infty$, le lecteur avisé montrera qu'en dérivant selon x on obtient :

$$f_Y(x) = (g^{-1})'(x) f_X\left(g^{-1}(x)\right) \iff f_Y(x) = \frac{1}{g'(g^{-1}(x))} f_X\left(g^{-1}(x)\right) \iff g'(g^{-1}(x)) f_Y(x) = f_X\left(g^{-1}(x)\right)$$

En posant enfin $t = g^{-1}(x)$, et ainsi $g(t) = x$ par le caractère bijectif de g dans le cas traité, (on peut utiliser $g' > 0$ aussi bien que $g' < 0$) ceci se réécrit $g'(t) = f_Y(g(t)) = f_X(t)$ et ce, pour tout t de $g^{-1}(I) = g^{-1}(]a;b[) =]\alpha;\beta[$. Le lecteur conclura en combinant les éléments établis.

Théorème de transfert (*formulation selon Gastineau*) : Soit f une densité d'une variable aléatoire X et φ une fonction définie sur $\mathbb{R}$. Sous réserve d'existence (si l'intégrale est absolument convergente) on a :

$$\mathbb{E}[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x)f(x)\mathrm{d}x$$

L'idée est, ici, de restreindre l'utilisation aux cas d'une fonction φ qui n'aurait pas de problème de domaine de définition pour la composer avec X. En particulier, on peut définir les moments d'ordre $r \in \mathbb{N}^*$ puisque $x \mapsto x^r$ est définie sur $\mathbb{R}$, tout comme les fonctions affines.

Exemple 2 : Soit X une variable aléatoire à densité f_X (donc définie sur $\mathbb{R}$, quitte à poser $f_X(x) = 0$ si $x \notin X(\Omega)$). Les fonctions $x \mapsto x^n$ sont définies et continues pour tout $n \in \mathbb{N}^*$ et on a alors, par le théorème de transfert :

$$m_r(X) = \mathbb{E}[X^r] = \int_{-\infty}^{+\infty} x^r f_X(x)\, \mathrm{d}x$$

sous couvert de convergence de cette dernière intégrale (en pratique, on la vérifiera). Ce résultat peut être retenu car souvent usité dans les problèmes en lien avec les probabilités à densité.

Théorème de transfert (*formulation expansive*) : Soit X une variable aléatoire à valeurs dans I, un intervalle, et g définie sur J contenant I, intégrable sur I. Si a et b sont les bornes de I (dans cet ordre) alors, sous couvert de convergence, on a :

$$\mathbb{E}[g(X)] = \int_a^b g(t)f_X(t)\, \mathrm{d}t$$

Cette formulation a l'avantage de regrouper les cas où l'intervalle I est un segment comme un ouvert afin d'éviter d'avoir à prouver la convergence dans ce cas, mais permet la souplesse de s'appliquer aussi au cas où l'intégrale serait impropre. La tournure *g est intégrable* est plus expansive que *g est continue sauf sur un nombre fini de points*, ce dernier cas ne regroupant pas toutes les possibilités d'utilisation.

Remarque : Si f_X est nulle hors du domaine $D \subset \mathbb{R}$ (D est alors le support de X), on pourrait écrire $f_X = f_X \cdot \mathbb{1}_D$. Avec la convention que, si $x \notin D$ alors la nullité de l'indicatrice $\mathbb{1}_D(x)$ est prioritaire sur l'évaluation (éventuellement impossible) de g en x, on pourrait réécrire :

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(t)\underbrace{f_X(t) \cdot \mathbb{1}_D(t)}\, \mathrm{d}t = \mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(t)f_X(t)\, \mathrm{d}t$$

sous couvert de convergence, et sans souci du domaine de définition réel de g. Cette pratique est $\boxed{\text{à éviter en ensD2}}$ (comme en filière EC) mais s'approche davantage de la formulation plus aboutie en mathématiques.
On peut lire par exemple, à cette fin, *Calcul des Probabilités* de D. Foata et A. Fuchs.