

Statistiques descriptives

Variable aléatoire quantitative discrète

Dans une population

Valeur x_i à la probabilité p_i

- **Moyenne théorique** $\mu = E(x) = \sum x_i \cdot p_i$
- **Variance théorique** $V = \sigma^2(X) = \sum (x_i - \mu)^2 \cdot p_i$
- **Ecart-type théorique** $\sigma = \sqrt{V}$

Variable aléatoire quantitative continue

Dans une population

$f(x)$ = densité de probabilité

- **Moyenne théorique** $\mu = E(x) = \int x \cdot f(x) dx$
- **Variance théorique** $V = \sigma^2(X) = \int (x - \mu)^2 \cdot f(x) dx$
- **Ecart-type théorique** $\sigma = \sqrt{V}$

Dans un échantillon (VA quantitatives discrète et continue)

Taille de l'échantillon N et ensemble de valeur x_i

- **Moyenne observée** $m = E(x) = \sum x_i / N$
- **Variance observée** $s^2 = \frac{1}{N} \sum (x_i - m)^2$
- **Ecart-type observé** $s = \sqrt{s^2}$

Cas de variable quantitative transformée en qualitative ordinale

À chaque classe est affectée une valeur (en général la valeur médiane de la classe)

Chaque sujet se voit affecter la valeur de la classe à laquelle il appartient

La moyenne des mesures pour l'échantillon de N sujets se calcule comme précédemment:

$$m = E(x) = \sum x_i / N$$

x_i étant la valeur de la classe à laquelle appartient le i-ème sujet

Propriétés des indices

Espérance (moyenne) : soient a et b deux constantes

$$E(X+a) = E(X) + a$$

$$E(aX) = a E(X)$$

Si X et Y sont 2 variables aléatoires

$$E(X + Y) = E(X) + E(Y)$$

$$E(X - Y) = E(X) - E(Y)$$

$$E(aX + bY) = a E(X) + b E(Y)$$

Variance

$$V(X + a) = V(X)$$

$$V(aX) = a^2 V(X)$$

Si X et Y sont 2 variables aléatoires indépendantes

$$V(X + Y) = V(X - Y) = V(X) + V(Y)$$

$$V(aX + bY) = a^2 V(X) + b^2 V(Y)$$

Autres indices

Médiane = valeur de X telle que

50% des individus ont $x <$ médiane

50% des individus ont $x >$ médiane

Premier quartile = valeur de X telle que

25% des individus ont $x <$ 1^{er} quartile

75% des individus ont $x >$ 1^{er} quartile

Troisième quartile = valeur de X telle que

75% des individus ont $x <$ 3^{ème} quartile

25% des individus ont $x >$ 3^{ème} quartile

Mode = Valeur de X de fréquence maximale

Etendue (range) : min ; max

Classe modale

Classe d'effectif le plus élevé

Le mode est alors la valeur médiane de la classe modale