Cours

Population : ensemble de référence sur lequel vont porter les observations.

Caractère : élément de mesure ; de deux types : qualitatif (H/F, couleur …) ou quantitatif (poids, taille...).

La variance v est la moyenne des carrés des écarts des valeurs de la série par rapport à la moyenne.

$$v = \sigma^2 = \sum_{i=1}^{n} f_i \left(\overline{x} - x_i\right)^2.$$

L'écart type évalue la dispersion des données d'une population en se basant sur un échantillon de cette population autour de la moyenne.

La loi binomiale : considérons *X*, une variable aléatoire discrète, associée à une expérience possédant l'alternative qu'un événement *E* soit réalisé (= succès de probabilité *p*) ou non réalisé ('non succès' de probabilité 1-p). On réalise *n* fois l'événement *E*.

Loi normale : *fct* de densité de probabilité : $$N(m,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

La loi χ^2 tend vers la loi normale centrée réduite quand $v \to \infty$; les deux lois sont quasi identiques quand $v > 30$.

On s'intéresse à la distribution du nombre de filles dans 100 familles de 4 enfants tirées au hasard dans une population.

Famille de x_i filles	Nombre de familles n_i	$n_i x_i$	$n_i x_i^2$
0	10	0	0
1	30	30	30
2	39	78	156
3	19	57	171
4	2	8	32
	$\Sigma\, n_i = 100$	$\Sigma\, n_i x_i = 173$	$\Sigma\, n_i x_i^2 = 389$

A partir des sommes discrètes, on peut facilement calculer moyenne pondérée m = 1,73 fille, variance v = 0,8971 fille^2 et écart type σ = 0,9472 fille.

Hypothèse : *loi binomiale* de moyenne : $m = np$ *avec* $m_{th} = m_{exp} = 1{,}73$ et $n = 4$.

Il vient que $p = m/n = 0{,}4325$ (proportion de filles) et $q = 1 - p = 0{,}5675$ (proportion de garçons).

On en déduit alors la probabilité théorique d'avoir k filles dans des familles de 4 enfants, et donnée par :

$$P_k = C_4^k \, (0{,}4325)^k \, (0{,}5675)^{4-k} \qquad \text{avec} \qquad 0 \leq k \leq 4.$$

Or les effectifs théoriques dans chaque catégorie k vaut $n_k = 100\, P_k$, avec $\Sigma\, n_k = \Sigma\, n_i = 100$; on peut donc reconstruire les distributions P_k et n_k.

Soit finalement :

Famille de x_i filles	Nombre de familles n_i	P_k	n_k
0	10	0,1037	10,37
1	30	0,3162	31,62
2	39	0,3615	36,15
3	19	0,1836	18,36
4	2	0,0350	3,50
	$\Sigma\, n_i = 100$	$\Sigma\, P_k = 1$	$\Sigma\, n_k = 100$