Centre Universitaire Abdelhafid-Boussouf –Mila-
Institut de science et technologie
Département de science et technologie

Cours analyse et modélisation hydrologique

Les tests d'adéquation d'une loi théorique

Master 01 : hydraulique urbaine

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- Supposons qu'une répartition empirique soit approchée par une courbe théorique F(x).

- Même si la courbe théorique est bien choisie, certains écarts entre celle-ci et les points expérimentaux sont inévitables.

- Une question se pose alors :

- Ces écarts sont-ils dus au hasard, vu le nombre limité d'observations?

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- Ou bien sont-ils structurels et proviennent-ils du fait que la courbe théorique a été mal choisie ?
- Pour y répondre, l'on fait appel aux tests d'adéquation ou de conformité.
- Dans ce cours, nous décrirons deux tests d'adéquation :
- Le test du Khi-deux χ^2,
- Le test de Kolmogorov Smirnov.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

* **L'application de ces derniers consiste à vérifier l'hypothèse H_0 selon laquelle une certaine loi de répartition F(x) s'ajuste à notre échantillon.**

- C'est à dire que H_0 est considérée comme hypothèse vraie et est appelée hypothèse nulle. Toute autre hypothèse est appelée hypothèse alternative H_1.
- Le risque consenti est choisi, et que nous appelons α de rejeter H_0 alors qu'elle est vraie, et est appelée seuil de signification. On a :

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- α = probabilité de rejeter H_0 alors qu'elle est vraie, oui bien,
- α = probabilité de nous tromper dans notre choix.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- **1-Test du χ^2 (Khi-deux)**

- Pour pouvoir faire des prévisions à l'aide d'un échantillon de données, on émet l'hypothèse H_o que cet échantillon appartient à une population dont les caractéristiques (moyenne et écart-type pour une loi normale, par exemple) sont égales à celles de l'échantillon).

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- Pour confirmer ou infirmer cette hypothèse, on utilise le test de Pearson, encore appelé le test du khi-deux (χ^2).
- Il permet de juger de la qualité de l'ajustement d'une distribution théorique à une distribution expérimentale.
- La procédure d'utilisation de ce test est la suivante :

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- 1- On divise l'intervalle de variation en k classes de façon que chacune d'elles contienne au minimum 5 données expérimentales (généralement, on s'abstient d'analyser des échantillons de moins de 10 valeurs).
- La classe i est bornée par les valeurs x_{i-1} et x_i choisies arbitrairement.
- On détermine la fréquence absolue observée ou l'effectif de chaque classe : f_{o1}, f_{o2}, …, f_{ok} avec $\sum f_{oi} = N$; N = taille de l'échantillon.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- 2- On peut obtenir les fréquences théoriques : $f_{t1} = N.p_1$, $f_{t2} = N.p_2$, … , $f_{tk} = N.p_k$ ou p_i représente la probabilité que la variable étudiée prenne une valeur appartenant à la classe i, c'est à dire que :

$$p_i = \int_{x_{i-1}}^{x_i} f(x)dx = FND_i - FND_{i-1}$$

- Ou f(x) = la fonction de densité de probabilité choisie, les limites de l'intégrale étant x_{i-1} et x_i.
- On a aussi $\sum f_{ti} = N$.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- 3- Pour évaluer l'ampleur de l'écart entre les fréquences absolues observées et les fréquences théoriques f_{ti} obtenues à partir de la loi théorique que l'on suppose adéquate, on utilise la quantité :

$$\chi^2 = \frac{\left(f_{o1} - f_{t1}\right)^2}{f_{t1}} + \frac{\left(f_{o2} - f_{t2}\right)^2}{f_{t2}} + \ldots + \frac{\left(f_{ok} - f_{tk}\right)^2}{f_{tk}} = \sum \frac{\left(f_{oi} - f_{ti}\right)^2}{f_{ti}}$$

- Pearson a demontré que la distribution de cette quantité est approximativement celle du khi-deux avce ν degrés de liberté; avec ν = k-1-r,

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- r = nombre de paramètres qui caractérisent complètement la distribution théorique (dans le cas de la loi Normale r = 2).
- Les conditions d'utilisation du test du χ^2 sont :
- a/ l'échantillon prélevé au hasard à partir de la population,
- b/ la taille de l'échantillon suffisamment importante.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

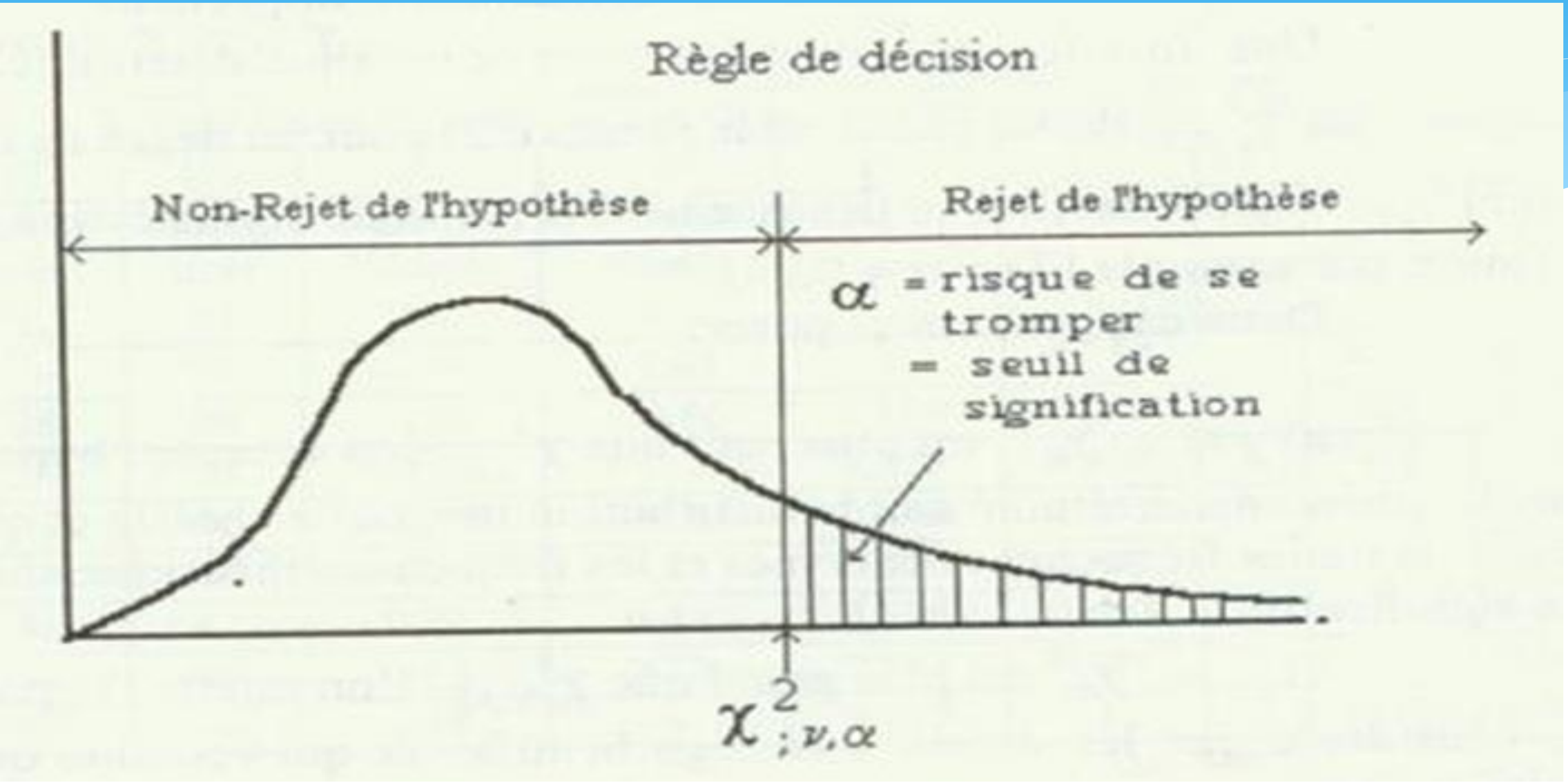


Figure 01 : Règle de la décision du test de KHI-DEUX

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- Une fois le χ_e^2 relatif à notre échantillon déterminé, on le compare au $\chi_{\nu,\alpha}^2$ donné par la table pour un degré de liberté connu ν et une probabilité au dépassement α (seuil de signification fixé à l'avance, par exemple, FD = α =0.05).
- Deux cas peuvent se poser :

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- a/ si χ_e^2 est plus petit que $\chi_{\nu,\alpha}^2$, l'on accepte l'hypothése que le phénoméne étudié suit la distribution théorique et que les écarts entre les fréquences observées et les fréquences théoriques ne sont pas significatifs.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- b/ si χ_e^2 est plus grand que $\chi_{\nu,\alpha}^2$, l'on rejette l'hypothèse H_0 considérée car les écarts sont significatifs; ce qui veut dire que les données expérimentales suivent une loi autre que celle de notre hypothèse, et l'on essaye une autre loi d'ajustement.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

Exemple : soit la série des pluies journalières maximales (Pjmax)

Année	pluie	Année	pluie	Année	pluie	Année	pluie	Année	pluie
1922	44	1933	35.4	1944	44.9	1955	27	1965	36.9
1923	29.7	1934	40.6	1945	21	1956	43.4	1966	24
1924	30.2	1935	26.6	1946	47.3	1957	42.6	1967	35
1925	40.5	1936	30	1947	39.7	1958	63.8	1968	27.5
1926	63	1937	40.5	1948	39.6	1959	32.4	1969	43.1
1927	35.1	1938	32.5	1949	29.9	1960	37.7	1970	48.4
1928	41.6	1939	31.2	1950	79.1	1961	35	1971	19.7
1929	49.5	1940	40.2	1951	28.7	1962	30.7	1972	37.5
1930	43.8	1941	45.8	1952	31	1963	28.2	1973	33.3
1931	53.5	1942	25.4	1953	30.6	1964	37.4	1974	41
1932	22.1	1943	40	1954	21.8				

Tableau 01 : *Série des pluies journalières maximales*

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

On essaye de savoir si une loi normale, avec une moyenne $\bar{P}$= 37.35 mm, et un écart-type S = 11.14mm, s'ajuste a notre échantillon de pluie journalières maximales (tableau 01) à un seuil de signification de 0.05. Pour cela on applique le Test de **KHI-DEUX** a notre échantillon .

Premièrement on calcule le nombre de classe par la formule suivante :

$K = \sqrt{N}$, avec N = 53.

$K \approx 8$.

Le tableau ci-dessous donne les détails de calculs :

- La colonne 01 donne la classe;
- Les colonnes 02 et 03 donnent respectivement les bornes inferieurs et supérieurs de chaque classe;
- Les colonnes 04 et 05 donnent respectivement la variable réduite correspondante à la limite inférieure et supérieure de chaque intervalle **$Z_i = (P_i - \bar{P})/S$**

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

* les colonnes 6 et 7 indiquent les probabilités au non dépassement (FNDi-1 et FNDi correspndant respectivement à la limite inférieure Xi-1 et la limite superieur Xi, la colonne 8 donne les fréquences expérimentales ou les effectifs Foi de chaque intervalle i; Foi est égale au nombre de valeurs qui se trouvent dans chaque intervalle i.

La colonne 9 donne les fréquences théoriques ou les effectifs théoriques Fti de chaque intervalle i, **(Fti = N (FNDi-FNDi-1)**, la colonne 10 indique le **Xi² : Xi² = (Foi-Fti)² /Fti**

Au bas de la colonne 10 nous avons $X^2 = \sum X_i{}^2$

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

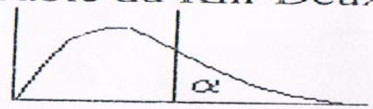
- On cherche maintenant le $X^2_{v\,;\alpha}$ avec $\alpha = 0.05$
- Et v = nombre de degrés de liberté = K-1-r ou
- K = nombre de classes = 8,
- r = 2 = nombre qui déterminent la loi normale.
- V= K-1-r => V = 8-1-2=> V = 5
- Le tableau du X^2 donne $X^2_{5\,;0.05} = 11.07$ (**voir la diapositive 20**)
- Comme le X^2 calculé est plus petit que celui donné par le tableau (8.5< 11.07) , on accepte donc l'hypothèse qu'une loi normale ayant une moyenne égale à 37.35 mm et un écart-type égal à 11.14 mm représente notre échantillon.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

$\alpha = 0.05$
$V = 5$

ANNEXE 2

Table du Khi-Deux



↓v, α→	0,995	0,99	0,975	0,95	0,75	0,5	0,25	0,1	0,05	0,025	0,01	0,005
1	4E-05	2E-04	1E-03	0,004	0,102	0,455	1,323	2,706	3,841	5,024	6,635	7,879
2	0,01	0,02	0,051	0,103	0,575	1,386	2,773	4,605	5,991	7,378	9,21	10,6
3	0,072	0,115	0,216	0,352	1,213	2,366	4,108	6,251	7,815	9,348	11,34	12,84
4	0,207	0,297	0,484	0,711	1,923	3,357	5,385	7,779	9,488	11,14	13,28	14,86
5	0,412	0,554	0,831	1,145	2,675	4,351	6,626	9,236	11,07	12,83	15,09	16,75
6	0,676	0,872	1,237	1,635	3,455	5,348	7,841	10,64	12,59	14,45	16,81	18,55
7	0,989	1,239	1,69	2,167	4,255	6,346	9,037	12,02	14,07	16,01	18,48	20,28
8	1,344	1,647	2,18	2,733	5,071	7,344	10,22	13,36	15,51	17,53	20,09	21,95
9	1,735	2,088	2,7	3,325	5,899	8,343	11,39	14,68	16,92	19,02	21,67	23,59
10	2,156	2,558	3,247	3,94	6,737	9,342	12,55	15,99	18,31	20,48	23,21	25,19
11	2,603	3,053	3,816	4,575	7,584	10,34	13,7	17,28	19,68	21,92	24,73	26,76
12	3,074	3,571	4,404	5,226	8,438	11,34	14,85	18,55	21,03	23,34	26,22	28,3
13	3,565	4,107	5,009	5,892	9,299	12,34	15,98	19,81	22,36	24,74	27,69	29,82
14	4,075	4,66	5,629	6,571	10,17	13,34	17,12	21,06	23,68	26,12	29,14	31,32
15	4,601	5,229	6,262	7,261	11,04	14,34	18,25	22,31	25	27,49	30,58	32,8
16	5,142	5,812	6,908	7,962	11,91	15,34	19,37	23,54	26,3	28,85	32	34,27
17	5,697	6,408	7,564	8,672	12,79	16,34	20,49	24,77	27,59	30,19	33,41	35,72
18	6,265	7,015	8,231	9,39	13,68	17,34	21,6	25,99	28,87	31,53	34,81	37,16
19	6,844	7,633	8,907	10,12	14,56	18,34	22,72	27,2	30,14	32,85	36,19	38,58
20	7,434	8,26	9,591	10,85	15,45	19,34	23,83	28,41	31,41	34,17	37,57	40
30	13,79	14,95	16,79	18,49	24,48	29,34	34,8	40,26	43,77	46,98	50,89	53,67
40	20,71	22,16	24,43	26,51	33,66	39,34	45,62	51,81	55,76	59,34	63,69	66,77
50	27,99	29,71	32,36	34,76	42,94	49,33	56,33	63,17	67,5	71,42	76,15	79,49
60	35,53	37,48	40,48	43,19	52,29	59,33	66,98	74,4	79,08	83,3	88,38	91,95
70	43,28	45,44	48,76	51,74	61,7	69,33	77,58	85,53	90,53	95,02	100,4	104,2
80	51,17	53,54	57,15	60,39	71,14	79,33	88,13	96,58	101,9	106,6	112,3	116,3
100	67,33	70,06	74,22	77,93	90,13	99,33	109,1	118,5	124,3	129,6	135,8	140,2

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

Solution :

Test de Khi-deux appliqué à la série de pluies journaliéres maximales									
1	2	3	4	5	6	7	8	9	10
N	Borne InF	Borne Sup	Variable réduite						
i	Pi-1	Pi	Zi-1	ZI	FNDi-1	FNDi	Foi	Fti	X^2
1	$-\infty$	26	$-\infty$	-1.02	0	0.154	6	8.15	0.57
2	26	30	-1.02	-0.66	0.154	0.255	7	5.30	0.55
3	30	33	-0.66	-0.39	0.255	0.348	8	4.90	1.96
4	33	36	-0.39	-0.12	0.348	0.452	5	5.70	0.09
5	36	40	-0.12	0.24	0.452	0.595	6	7.50	0.3
6	40	42	0.24	0.42	0.595	0.663	7	3.60	3.21
7	42	46	0.42	0.78	0.663	0.782	7	6.3	0.03
8	46	$+\infty$	0.78	$+\infty$	0.782	1	7	11.54	1.79
								$\sum X^2 = 8.50$	

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- **2-Test de Kolmogorov-Smirnov**
- C'est un test d'ajustement qui permet de comparer une distribution de valeurs observées à une distribution théorique. Ce test joue le même rôle que celui du khi-deux.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- Pour accepter ou rejeter l'hypothèse H_0 que la loi choisie s'ajuste bien à notre échantillon, on considère la grandeur D qui caractérise la différence entre la répartition empirique et la répartition théorique.
- La grandeur D peut être choisie de plusieurs façons.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- Cette grandeur D est elle-même une variable aléatoire dont la loi de répartition, dans certains cas, pour N suffisamment grand, ne dépend pratiquement pas de la fonction F(x).
- Avec le test de Kolmogorov-Smirnov, on cherche la valeur maximale de la valeur absolue de la différence entre la fonction de répartition empirique $F_N(x)$ d'un échantillon de N valeurs et la fonction de répartition théorique F(x) correpondante soit :

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

$$D_N = D_{\max} = \max\left|F_N(x) - F(x)\right|$$

- Kolmogorov a montré que, quelque soit la fonction de répartition F(x) d'une variable continue x, lorsque le nombre d'observations augmente, la fonction de répartition de la grandeur $D_N*(N)^{1/2}$ tend asymptotiquement vers

$$\Pr ob\left(D_N \sqrt{N}\right) \rightarrow K(y) = \sum_{-\infty}^{+\infty} (-1)^k \exp\left(-2k^2 y^2\right)$$

- Les valeurs de cette probabilité ont été tabulées.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

- On rejettera l'hypothèse H_o que la loi choisie représente notre échantillon, au niveau de signification choisi, lorsque D_N est supérieur ou égal à d_n, qui est la valeur de l'écart théorique.

LES TESTS D'ADEQUATION D'UNE LOI THEORIQUE

N = 20;
α = 0.05

ANNEXE 3

Table du test de Kolmogorov-Smirnov
$D_n = \mathrm{Sup} \mid F_n^*(x) - f(x) \mid$
Valeurs de d_n telles que $P = P(D_n < d_n)$

n	P=.80	P=.90	P=.95	P=.98	P=.99
1	.90000	.95000	.97500	.99000	.99500
2	.68377	.77639	.84189	.90000	.92929
3	56481	.63604	.70760	.78456	.82900
4	.49265	.56522	.62394	.68887	.73424
5	.44698	.50945	.56328	.62718	.66853
6	.41037	.46799	.51926	.57741	.61661
7	.38148	.43607	.48342	.53844	.57581
8	.35381	.40962	.45427	.51654	.54179
9	.33910	.38746	.43001	.47960	.51332
10	.32260	.36866	.40925	.45662	.48893
11	.30829	.35242	.39122	.43670	.46770
12	.29577	.33815	.37543	.41918	.44905
13	.28470	.32549	.36143	.40362	.43247
14	.27481	.31417	.34890	.38970	.41762
15	.26588	.30397	.33760	.37713	.40420
16	.25778	.29472	.32733	.36571	.39201
17	.25039	.28627	.31796	.35528	.38086
18	.24360	.27851	.30936	.34569	.37062
19	.23735	.27136	.30143	.33685	.36117
20	.23156	.26473	29408	.32866	.35241
21	.22617	.25858	.28724	.32104	34427
22	.22115	.25283	.28087	.31394	.33666
23	.21645	.24746	.27490	.30728	.32954
24	.21205	.24242	.26931	.30104	.32286
25	.20790	.23768	.26404	.29516	.31657
26	.20399	.23320	.25907	.28962	.31064
27	.20030	.22898	.25438	.28438	.30502
28	.19680	.22497	.24993	.27942	.29971
29	.19348	.22117	.24571	.27471	.29466
30	.19032	.21756	.24170	.27023	.28987
31	.18732	.21412	.23788	.26596	28530
32	.18445	.21085	.23424	.26189	.28094
33	.18171	.20771	.23076	.25801	.27677
34	.17909	.20472	.22743	.25429	.27279
35	.17659	.20185	.22425	.25073	.26897