

TD 8 : tests du  $\chi^2$ 

Les questions marquées d'un astérisque (\*) sont facultatives.

Dans tous les exercices, on détaillera avec autant de soin que possible toutes les étapes des tests mis en oeuvre.

Voir en fin de feuille pour des quantiles usuels de loi du  $\chi^2$  (table produite par Claude Bélisle).

**Exercice 1. (Questions de cours)** Rappeler les hypothèses, statistique de test et loi de la statistique de test sous  $H_0$  des tests du chi-deux d'ajustement et d'indépendance.

**Exercice 2.** On s'intéresse à la production d'un anticorps par les individus en fonction de leur couleur de cheveux. Après étude, on constate que 10% des Français produisent naturellement cet anticorps. On recrute 50 personnes aux cheveux blonds, et on constate que 10 d'entre eux le produisent.

Peut-on dire que la couleur de cheveux a un impact significatif, avec un niveau 5%, sur la production de cet anticorps ?

**Exercice 3.** On étudie l'impact des séjours dans l'espace sur le génôme des bactéries. Pour cela, on considère un gène admettant trois mutations A, B et C. Sur Terre, les proportions entre ces mutations sont de 25% pour la mutation A, 40% pour la mutation B et 35% pour la mutation C. Sur 66 échantillons cultivés dans l'espace, on observe 12 mutations A, 33 mutations B et 21 mutations C.

Peut-on dire que le séjour dans l'espace a impacté de façon significative, avec un niveau 5%, les mutations des bactéries ?

**Exercice 4.** Un organisme de cours à domicile teste trois méthodes d'entraînement pour un concours. On observe les résultats d'admission suivants :

	Admis	Recalé
Méthode 1	19	25
Méthode 2	41	35
Méthode 3	17	33

Y a-t-il une différence significative, avec un niveau 5%, entre les méthodes ?

**Exercice 5.** On échantillonne des population d'insectes de la même espèce dans trois environnements différents : le premier est un champs cultivé, le deuxième une forêt, et le troisième des bosquets isolés. Pour chaque environnement, on mesure la concentration d'un marqueur biologique relié à la pollution. La mesure fournit un réel positif, qu'on résume comme suit. Chaque case contient le nombre d'insectes observés ayant une concentration de marqueur dans l'intervalle précisé.

	$[0, 20[$	$[20, 30[$	$[30 - 40[$	$[40, +\infty[$
Champs	5	11	37	14
Forêt	8	23	27	5
Bosquets	10	19	30	9

Peut-on dire, avec un test de niveau 5%, que l'impact de la pollution sur les insectes est indépendant de l'environnement ?

LOI DU KHI-DEUX AVEC  $k$  DEGRÉS DE LIBERTÉ  
QUANTILES D'ORDRE  $1 - \gamma$

$k$	$\gamma$										
	0.995	0.990	0.975	0.950	0.900	0.500	0.100	0.050	0.025	0.010	0.005
1	0.00	0.00	0.00	0.00	0.02	0.45	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	0.21	1.39	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	2.37	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	3.36	7.78	9.94	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.81	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.28	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	79.33	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	89.33	107.57	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17

Si  $k$  est entre 30 et 100 mais n'est pas un multiple de 10, on utilise la table ci-haut et on fait une interpolation linéaire. Si  $k > 100$  on peut, grâce au théorème limite central, approximer la loi  $\chi^2(k)$  par la loi  $N(k, 2k)$ .