Partiel de Statistique

Mercredi 14 Janvier 2004, 8h00-10h00

On appelle loi de Pareto de paramètres $a > 0$ et $\lambda > 0$ une loi de fonction de répartition

$$F(x) = \begin{cases} 1 - \left(\frac{a}{x}\right)^{\lambda} & \text{si } x > a \\ 0 & \text{sinon} \end{cases}$$

On utilisera la notation habituelle $X \sim P_{a,\lambda}$ pour désigner une variable aléatoire X distribuée suivant la loi de Pareto de paramètres a et λ.

Partie 1 : Estimation de λ

On suppose dans cette partie que le paramètre a est connu. On s'intéresse alors à l'estimation du paramètre λ à partir de n observations $x_1, ..., x_n$ issues d'un échantillon $(X_1, ..., X_n)$ distribué suivant la loi de Pareto $P_{a,\lambda}$.

1) Etudier avec soin la vraisemblance des observations et en déduire l'estimateur du maximum de vraisemblance du paramètre λ noté $\hat{\lambda}_{MV}$.

2) Déterminer la borne de Cramer-Rao associée à un estimateur non biaisé du paramètre λ.

3) On désire étudier les performances de l'estimateur $\hat{\lambda}_{MV}$ (*Remarque : si les questions 3.1) et 3.2) vous posent des problèmes, vous pouvez admettre que* $\hat{\lambda}_{MV}$ *s'écrit* $\hat{\lambda}_{MV} = \frac{n}{V}$ *avec* $V \sim \Gamma(\lambda, n)$ *et passez à la question 3.3).*

3.1) On pose $V_i = \ln\left(\frac{X_i}{a}\right)$. Montrer que V_i suit une loi Gamma dont on précisera les paramètres.

3.2) Soient X et Y deux variables aléatoires indépendantes de lois Gamma $\Gamma(\alpha, \beta_1)$ et $\Gamma(\alpha, \beta_2)$. Montrer que $X + Y$ suit une loi Gamma $\Gamma(\alpha, \beta_1 + \beta_2)$. En déduire que l'estimateur $\hat{\lambda}_{MV}$ s'écrit $\hat{\lambda}_{MV} = \frac{n}{V}$, où V est une variable aléatoire (que l'on précisera) distribuée suivant une loi Gamma de paramètres λ et n.

3.3) On rappelle le résultat suivant

$$\int_0^{+\infty} u^n e^{-u}\, du = \Gamma(n+1) = n!$$

Etudier la moyenne et la variance de $\hat{\lambda}_{MV}$ et en déduire un estimateur sans biais et convergent du paramètre $\hat{\lambda}_{MV}$. Est-il efficace ?

4) On suppose dans cette question qu'on dispose d'informations a priori sur le paramètre λ résumées dans la loi Gamma $\Gamma(c, d)$ (avec $c > 0$ et $d > 1$)de densité

$$f(\lambda) \propto e^{-c\lambda} \lambda^{d-1} I_{\mathbb{R}^+}(\lambda)$$

Déterminer l'estimateur du maximum a posteriori du paramètre λ noté $\hat{\lambda}_{MAP}$ et analyser le comportement de cet estimateur pour n "petit" et "grand".

Partie 2 : Test d'hypothèses

On dispose de n observations $x_1, ..., x_n$ associées à un échantillon $(X_1, ..., X_n)$ de loi de Pareto $P_{a,\lambda}$. On suppose à nouveau que le paramètre a est connu et on désire étudier le test d'hypothèses suivant :

$$\left\{ \begin{array}{ll} H_0 : & \lambda = \lambda_0 \\ H_1 : & \lambda = \lambda_1 \end{array} \right.$$

avec $\lambda_1 > \lambda_0 > 0$.

1) Expliciter le test de Neyman-Pearson associé à ce problème et montrer qu'il consiste à comparer la statistique

$$V = \sum_{i=1}^{n} \ln\left(\frac{X_i}{a}\right)$$

à un seuil noté K_α. On rappelle que conformément à la partie précédente, V_i est distribuée suivant une loi Gamma de paramètres λ et 1, i.e. $V_i \sim \Gamma(\lambda, 1)$. Calculer la moyenne de $V_i = \ln\left(\frac{X_i}{a}\right)$ et commenter la forme du test de Neyman-Pearson.

2) On rappelle que d'après les résultats de la partie 1, on a $V \sim \Gamma(\lambda, n)$. On suppose de plus qu'on peut calculer la fonction Q_n définie par

$$Q_n(x) = \frac{1}{\Gamma(n)} \int_0^x u^{n-1} e^{-u}\, du$$

ainsi que son inverse $Q_n^{-1}(x)$ pour toutes les valeurs de n et de x. Déterminer la valeur du seuil K_α en fonction de λ_0, α et de Q_n.

3) Exprimer la puissance de ce test en fonction de λ_0, λ_1 et α (en utilisant les fonctions Q_n et Q_n^{-1}). Déterminer le rapport signal sur bruit du test en fonction de λ_0 et λ_1 (on suppose que n et α sont fixés). En déduire lequel des deux test suivants est le plus puissant :

- Test 1 : $\lambda_0 = 1$ contre $\lambda_1 = 2$
- Test 2 : $\lambda_0 = 10$ contre $\lambda_1 = 20$

4) Lorsqu'on ne peut calculer les fonctions Q_n et Q_n^{-1}, on peut approcher la loi de V par une loi normale d'après le théorème de la limite centrale. Calculer le seuil K_α et la puissance du test lorsqu'on utilise cette loi approchée et que les paramètres du test sont $\alpha = 0.01$, $\lambda_0 = 1, \lambda_1 = 2$ et $n = 40$.

5) Les observations $x_1, ..., x_{20}$ sont réunies dans le tableau ci-dessous

1.05	2.46	1.50	1.72	1.12	1.27	1.78	4.99	1.20	1.81
1.49	1.23	1.08	1.30	2.74	1.90	1.07	1.90	1.11	3.85

On désire effectuer un test du chi-deux pour déterminer si ces observations sont issues d'une loi de Pareto $P_{1,1}$ de densité $f(x) = \frac{1}{x^2} I_{]1,+\infty[}(x)$. Découper le domaine de définition de cette densité en $K = 4$ classes équiprobables. Que conclure avec les risques $\alpha = 0.01$ et les risques $\alpha = 0.05$?

LOI	**Densité de probabilité**	**m**	$\boldsymbol{\sigma}^2$	**F. C.**
Uniforme	$f(x) = \frac{1}{b-a}$ $x \in]a, b[$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\dfrac{e^{itb} - e^{ita}}{it(b-a)}$
Gamma $\Gamma(\theta, \nu)$	$f(x) = \frac{\theta^\nu}{\Gamma(\nu)} e^{-\theta x} x^{\nu-1}$ $\theta > 0,\ \nu > 0$ $x \geq 0$	$\dfrac{\nu}{\theta}$	$\dfrac{\nu}{\theta^2}$	$\dfrac{1}{\left(1 - i\frac{t}{\theta}\right)^\nu}$
Première loi de Laplace	$f(x) = \frac{1}{2} e^{-\lvert x \rvert}$	0	2	$\dfrac{1}{1+t^2}$
Normale $N(m, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$	m	σ^2	$e^{imt - \frac{\sigma^2 t^2}{2}}$
Khi$_2$ χ^2_ν $\Gamma\left(\frac{1}{2}, \frac{\nu}{2}\right)$	$f(x) = k e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}$ $k = \dfrac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}$ $\nu \in \mathbb{N}^*,\ x \geq 0$	ν	2ν	$\dfrac{1}{(1-2it)^{\frac{\nu}{2}}}$
Cauchy $c_{\lambda,\alpha}$	$f(x) = \dfrac{1}{\pi\lambda\left(1 + \left(\frac{x-\alpha}{\lambda}\right)^2\right)}$ $\lambda > 0,\ \alpha \in \mathbb{R}$	(-)	(-)	$e^{i\alpha t - \lambda \lvert t \rvert}$
Bêta	$f(x) = k x^{a-1} (1-x)^{b-1}$ $k = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$ $a > 0,\ b > 0,\ x \in]0, 1[$	$\frac{a}{a+b}$	$\frac{ab}{(a+b)^2(a+b+1)}$	(*)

Loi Normale $N(0,1)$

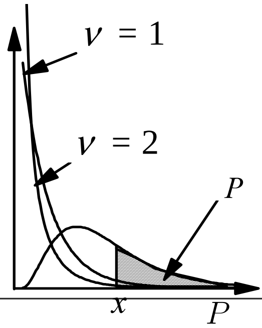
$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{u^2}{2}}\, du$$

$$\Phi(-x) = 1 - \Phi(x)$$

x	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	0.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	0.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	0.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	0.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	0.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	0.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	0.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	0.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	0.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	0.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	0.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	0.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	0.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	0.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	0.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	0.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	0.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	0.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	0.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	0.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	0.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	0.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	0.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	0.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	0.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	0.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	0.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	0.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	0.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	0.9985	.9986	.9986

Distribution du χ^2

$\boxed{P\,[\chi^2_\nu \geq x] = P}$



	P						
ν	**0.99**	**0.975**	**0.95**	**0.90**	**0.80**	**0.70**	**0.60**
1	0.000	0.001	0.004	0.016	0.064	0.148	0.275
2	0.020	0.051	0.103	0.211	0.446	0.713	1.022
3	0.115	0.216	0.352	0.584	1.005	1.424	1.869
4	0.297	0.484	0.711	1.064	1.649	2.195	2.753
5	0.554	0.831	1.145	1.610	2.343	3.000	3.655
6	0.872	1.237	1.635	2.204	3.070	3.828	4.570
7	1.239	1.690	2.167	2.833	3.822	4.671	5.493
8	1.646	2.180	2.733	3.490	4.594	5.527	6.423
9	2.088	2.700	3.325	4.168	5.380	6.393	7.357
10	2.558	3.247	3.940	4.865	6.179	7.267	8.295
11	3.053	3.816	4.575	5.578	6.989	8.148	9.237
12	3.571	4.404	5.226	6.304	7.807	9.034	10.182
13	4.107	5.009	5.892	7.042	8.634	9.926	11.129
14	4.660	5.629	6.571	7.790	9.467	10.821	12.078
15	5.229	6.262	7.261	8.547	10.307	11.721	13.030
16	5.812	6.908	7.962	9.312	11.152	12.624	13.983
17	6.408	7.564	8.672	10.085	12.002	13.531	14.937
18	7.015	8.231	9.390	10.865	12.857	14.440	15.893
19	7.633	8.907	10.117	11.651	13.716	15.352	16.850
20	8.260	9.591	10.851	12.443	14.578	16.266	17.809
21	8.897	10.283	11.591	13.240	15.445	17.182	18.768
22	9.542	10.982	12.338	14.041	16.314	18.101	19.729
23	10.196	11.689	13.091	14.848	17.187	19.021	20.690
24	10.856	12.401	13.848	15.659	18.062	19.943	21.652
25	11.524	13.120	14.611	16.473	18.940	20.867	22.616
26	12.198	13.844	15.379	17.292	19.820	21.792	23.579
27	12.879	14.573	16.151	18.114	20.703	22.719	24.544
28	13.565	15.308	16.928	18.939	21.588	23.647	25.509
29	14.256	16.047	17.708	19.768	22.475	24.577	26.475
30	14.953	16.791	18.493	20.599	23.364	25.508	27.442
40	22.164	24.433	26.509	29.051	32.345	34.872	37.134
50	29.707	32.357	34.764	37.689	41.449	44.313	46.864
60	37.485	40.482	43.188	46.459	50.641	53.809	56.620

Distribution du χ^2(suite)

	P							
ν	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.025**	**0.01**
1	0.455	0.708	1.074	1.642	2.706	3.841	5.024	6.635
2	1.386	1.833	2.408	3.219	4.605	5.991	7.378	9.210
3	2.366	2.946	3.665	4.642	6.251	7.815	9.348	11.345
4	3.357	4.045	4.878	5.989	7.779	9.488	11.143	13.277
5	4.351	5.132	6.064	7.289	9.236	11.071	12.833	15.086
6	5.348	6.211	7.231	8.558	10.645	12.592	14.449	16.812
7	6.346	7.283	8.383	9.803	12.017	14.067	16.013	18.475
8	7.344	8.351	9.524	11.030	13.362	15.507	17.535	20.090
9	8.343	9.414	10.656	12.242	14.684	16.919	19.023	21.666
10	9.342	10.473	11.781	13.442	15.987	18.307	20.483	23.209
11	10.341	11.530	12.899	14.631	17.275	19.675	21.920	24.725
12	11.340	12.584	14.011	15.812	18.549	21.026	23.337	26.217
13	12.340	13.636	15.119	16.985	19.812	22.362	24.736	27.688
14	13.339	14.685	16.222	18.151	21.064	23.685	26.119	29.141
15	14.339	15.733	17.322	19.311	22.307	24.996	27.488	30.578
16	15.338	16.780	18.418	20.465	23.542	26.296	28.845	32.000
17	16.338	17.824	19.511	21.615	24.769	27.587	30.191	33.409
18	17.338	18.868	20.601	22.760	25.989	28.869	31.526	34.805
19	18.338	19.910	21.689	23.900	27.204	30.144	32.852	36.191
20	19.337	20.951	22.775	25.038	28.412	31.410	34.170	37.566
21	20.337	21.991	23.858	26.171	29.615	32.671	35.479	38.932
22	21.337	23.031	24.939	27.301	30.813	33.924	36.781	40.289
23	22.337	24.069	26.018	28.429	32.007	35.172	38.076	41.638
24	23.337	25.106	27.096	29.553	33.196	36.415	39.364	42.980
25	24.337	26.143	28.172	30.675	34.382	37.652	40.646	44.314
26	25.336	27.179	29.246	31.795	35.563	38.885	41.923	45.642
27	26.336	28.214	30.319	32.912	36.741	40.113	43.195	46.963
28	27.336	29.249	31.391	34.027	37.916	41.337	44.461	48.278
29	28.336	30.283	32.461	35.139	39.087	42.557	45.722	49.588
30	29.336	31.316	33.530	36.250	40.256	43.773	46.979	50.892
40	39.335	41.622	44.165	47.269	51.805	55.758	59.342	63.691
50	49.335	51.892	54.723	58.164	63.167	67.505	71.420	76.154
60	59.335	62.135	65.227	68.972	74.397	79.082	83.298	88.379

Pour $30 \leq \nu \leq 100$, $\sqrt{2\chi^2_\nu} - \sqrt{2\nu - 1}$ suit approximativement une loi $N(0,1)$.

Pour $\nu > 100$, $(\chi^2_\nu - \nu)/\sqrt{2\nu}$ suit approximativement une loi $N(0,1)$.