

Chapitre 5 : Statistique inférentielle/tests

I- Loïs usuelles et non usuelles

Lois discrètes	Fonctions R	Espérance Variance	Fonction de masse $P(X = x)$
Binomiale(m, α)	<code>dbinom(x,size=m,prob=α)</code> <code>pbinom(q,size=m,prob=α)</code> <code>qbinom(p,size=m,prob=α)</code> <code>rbinom(n,size=m,prob=α)</code>	$m\alpha$ $m\alpha(1 - \alpha)$	$\binom{m}{x}\alpha^x(1 - \alpha)^{m-x}$
Poisson(λ)	<code>dpois(x,lambda=λ)</code> <code>ppois(q,lambda=λ)</code> <code>qpois(p,lambda=λ)</code> <code>rpois(n,lambda=λ)</code>	λ λ	$e^{-\lambda} \frac{\lambda^x}{x!}$
Géométrique(α)	<code>dgeom(x,prob=α)</code> <code>pgeom(q,prob=α)</code> <code>qgeom(p,prob=α)</code> <code>rgeom(n,prob=α)</code>	$\frac{1}{\alpha}$ $\frac{1-\alpha}{\alpha^2}$	$(1 - \alpha)^{x-1}\alpha$
Hyper- géométrique(m, n, k)	<code>dhyper(x,m=m,n=n,k=k)</code> <code>phyper(q,m=m,n=n,k=k)</code> <code>qhyper(p,m=m,n=n,k=k)</code> <code>rhyper(nn,m=m,n=n,k=k)</code>	$\frac{nm}{N}$ (avec $N = n + m$) $\frac{n(m/N)(1-(m/N))(N-n)}{(N-1)}$	$\frac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}$
Binomiale négative(m, α)	<code>dnbinom(x,size=m,prob=α)</code> <code>pnbinom(q,size=m,prob=α)</code> <code>qnbinom(p,size=m,prob=α)</code> <code>rnbinom(n,size=m,prob=α)</code>	$m \frac{1-\alpha}{\alpha}$ $m \frac{1-\alpha}{\alpha^2}$	$\binom{x+m-1}{m-1}\alpha^m(1 - \alpha)^x$
Uniforme discrète $\{1, \dots, m\}$	<code>(x %in% 1:m)/m</code> <code>sum(1:m<=q)/m</code> <code>match(1,1:m/m>=p)</code> <code>sample(x=1:m,size=n,TRUE)</code>	$\frac{m+1}{2}$ $\frac{m^2-1}{12}$	$\frac{1}{m}\mathbf{1}_{\{1,\dots,m\}}(x)$

Lois continues	Fonctions R	Espérance Variance	Densité
Normale(μ, σ^2)	dnorm(x, mean= μ , sd= σ) pnorm(q, mean= μ , sd= σ) qnorm(p, mean= μ , sd= σ) rnorm(n, mean= μ , sd= σ)	μ σ^2	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
Student(ν, μ)	dt(x, df= ν , ncp= μ) pt(q, df= ν , ncp= μ) qt(p, df= ν , ncp= μ) rt(n, df= ν , ncp= μ)	$\mu \sqrt{\frac{\nu}{2}} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}$ ($\nu > 1$) $\frac{\nu(1+\mu^2)}{\nu-2} - \frac{\mu^2\nu}{2} \times$ $\left(\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\right)$, ($\nu > 2$)	$\frac{\gamma^{\nu/2} e^{-\gamma\mu^2/2(x^2+\gamma)}}{\sqrt{\pi}\Gamma(\nu/2)2^{(\nu-1)/2}(x^2+\gamma)^{(\nu+1)/2}}$ $\times \int_0^\infty t^\nu e^{\frac{-\mu x}{2\sqrt{x^2+\gamma}} - \frac{t}{2}} dt$
Khi-deux(k, λ)	dchisq(x, df= k , ncp= λ) pchisq(q, df= k , ncp= λ) qchisq(p, df= k , ncp= λ) rchisq(n, df= k , ncp= λ)	$k + \lambda$ $2(k + 2\lambda)$	$\frac{1}{2} e^{-(x+\lambda)/2} \left(\frac{x}{\lambda}\right)^{k/4-1/2}$ $\times I_{k/2-1}(\sqrt{\lambda x})$
Fisher(ν_1, ν_2, λ)	df(x, df1= ν_1 , df2= ν_2 , ncp= λ) pf(q, df1= ν_1 , df2= ν_2 , ncp= λ) qf(p, df1= ν_1 , df2= ν_2 , ncp= λ) rf(n, df1= ν_1 , df2= ν_2 , ncp= λ)	$\frac{\nu_2(\nu_1+\lambda)}{\nu_1(\nu_2-2)}$ ($\nu_2 > 2$) $2 \frac{(\nu_1+\lambda)^2 + (\nu_1+2\lambda)(\nu_2-2)}{(\nu_2-2)^2(\nu_2-4)}$ $\times \left(\frac{\nu_2}{\nu_1}\right)^2$, ($\nu_2 > 4$)	$\sum_{k=0}^\infty \frac{e^{-\lambda/2} (\lambda/2)^k}{B(\frac{\nu_2}{2}, \frac{\nu_1}{2} + k) k!} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2} + k}$ $\times \left(\frac{\nu_2}{\nu_2 + \nu_1 x}\right)^{\frac{\nu_1 + \nu_2}{2} + k} x^{\frac{\nu_1}{2} - 1 + k}$
Exponentielle(λ)	dexp(x, rate= λ) pexp(q, rate= λ) qexp(p, rate= λ) rexp(n, rate= λ)	$\frac{1}{\lambda}$ $\frac{1}{\lambda^2}$	$\lambda e^{-\lambda x} \mathbf{1}_{\{x \geq 0\}}$
Uniforme(a, b)	dunif(x, min= a , max= b) punif(q, min= a , max= b) qunif(p, min= a , max= b) runif(n, min= a , max= b)	$\frac{a+b}{2}$ $\frac{(b-a)^2}{12}$	$\frac{1}{b-a} \mathbf{1}_{\{a \leq x \leq b\}}$
Bêta(α, β, λ)	dbeta(x, shape1= α , shape2= β , ncp= λ) pbeta(q, shape1= α , shape2= β , ncp= λ) qbeta(p, shape1= α , shape2= β , ncp= λ) rbeta(n, shape1= α , shape2= β , ncp= λ)	$\approx 1 - \frac{\beta}{C} \left(1 + \frac{\lambda}{2C^2}\right)$ avec $C = \alpha + \beta + \frac{\lambda}{2}$ $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ si $\lambda = 0$	$\sum_{i=0}^\infty P(i; \frac{\lambda}{2}) I'_x(\alpha + i, \beta)$
Cauchy(x_0, γ)	dcauchy(x, location= x_0 , scale= γ) pcauchy(q, location= x_0 , scale= γ) qcauchy(p, location= x_0 , scale= γ) rcauchy(n, location= x_0 , scale= γ)	Non définie Non définie	$\frac{1}{\pi} \left[\frac{\gamma}{(x-x_0)^2 + \gamma^2} \right]$
Logistique(μ, s)	dlogis(x, location= μ , scale= s) plogis(q, location= μ , scale= s) qlogis(p, location= μ , scale= s) rlogis(n, location= μ , scale= s)	μ $\frac{\pi^2}{3} s^2$	$\frac{1}{4s} \operatorname{sech}^2\left(\frac{x-\mu}{2s}\right)$
Log-Normale(μ, σ)	dlnorm(x, meanlog= μ , sdlog= σ) plnorm(q, meanlog= μ , sdlog= σ) qlnorm(p, meanlog= μ , sdlog= σ) rlnorm(n, meanlog= μ , sdlog= σ)	$e^{\mu+\sigma^2/2}$ $(e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}$
Gamma(α, β)	dgamma(x, shape= α , rate= β) pgamma(q, shape= α , rate= β) qgamma(p, shape= α , rate= β) rgamma(n, shape= α , rate= β)	$\alpha\beta$ $\alpha\beta^2$	$x^{\alpha-1} \frac{e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} \mathbf{1}_{x>0}$
Weibull(λ, k)	dweibull(x, shape= λ , scale= k) pweibull(q, shape= λ , scale= k) qweibull(p, shape= λ , scale= k) rweibull(n, shape= λ , scale= k)	$\lambda \Gamma\left(1 + \frac{1}{k}\right)$ $\lambda^2 \Gamma\left(1 + \frac{2}{k} - \mu^2\right)$	$\frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} \mathbf{1}_{x \geq 0}$
Gumbel(μ, β) (Package evd)	dgumbel(x, loc= μ , scale= β) pgumbel(q, loc= μ , scale= β) qgumbel(p, loc= μ , scale= β) rgumbel(n, loc= μ , scale= β)	$\mu + \beta$ $\frac{\pi^2}{6} \beta^2$	$\frac{ze^{-z}}{\beta}$ avec $z = e^{-\frac{x-\mu}{\beta}}$

II - Intervalles de confiance et tests d'hypothèses

Paramètre	Notation	Estimateur	Estimation	Fonction R
moyenne	μ	\bar{X}	\bar{x}	mean()
variance	σ^2	$\hat{\sigma}^2$	$\hat{\sigma}^2$	var()
médiane	m_e	\widehat{m}_e	\widehat{m}_e	median()
corrélation	ρ	$\hat{\rho}$	$\hat{\rho}$	cor()
proportion	p	\hat{p}	\hat{p}	mean()

Loi	Notation	Fonction R
Normale : $\mathcal{N}(0, 1)$	u_p	qnorm(p)
Student à n d.d.l. : $\mathcal{T}(n)$	t_p^n	qt($p, df=n$)
Khi-deux à n d.d.l. : $\chi^2(n)$	q_p^n	qchisq($p, df=n$)
Fisher à n et m d.d.l. : $\mathcal{F}(n, m)$	$f_p^{n,m}$	qf($p, df1=n, df2=m$)
d.d.l. : degrés de liberté		

Résumé sur les intervalles de confiance.

Type	Condition de validité	Fonction R
proportion	$np \geq 5$ et $n(1-p) \geq 5$ aucune	prop.test(x)\$conf binom.test(x)\$conf
moyenne	$n > 30$ ou normalité	t.test(x)\$conf
variance	normalité	sigma2.test(x)\$conf
médiane	aucune	wilcox.test(x)\$conf
corrélation	binormale	cor.test(x)\$conf

III- Tests paramétriques

Nature	Données	Conditions de validité	Fonction R
Tests paramétriques :			
moyenne	1 échantillon	$n > 30$ ou normalité	<code>t.test(x,...)</code>
	2 échantillons	normalité et variances égales	<code>t.test(x,y,...)</code>
	2 échantillons		<code>t.test(x,y,var.equal=F)</code>
	2 éch. appariés	$n > 30$ ou normalité	<code>t.test(x,y,paired=T)</code>
variance	1 échantillon	normalité	<code>sigma2.test(x,...)</code>
	2 échantillons	normalité	<code>var.test(x,y,...)</code>
	2 échantillons	grand échantillon	<code>asympt.test(x,y,...)</code>
corrélation	1 échantillon	normalité, $\mathcal{H}_0 : \rho = \rho_0$	<code>cor.test(x,y..)</code>
	2 échantillons	normalité	<code>cor.test.2.sample(x,y,...)</code>
proportion	1 échantillon	$np \geq 5$ et $n(1 - p) \geq 5$	<code>prop.test(x,...)</code>
	1 échantillon		<code>binom.test(x,...)</code>
	2 échantillons	grand échantillon	<code>prop.test(x,y,...)</code>
Tests d'indépendance :			
χ^2 d'indépendance	tableau de contingence	effectifs théoriques ≥ 5	<code>chisq.test(,correct=F)</code>
χ^2 de Yates	tableau 2×2	effectifs théoriques ≥ 2.5	<code>chisq.test()</code>
Fisher exact	tableau de contingence		<code>fisher.test()</code>
Tests d'adéquation :			
Shapiro-Wilk	1 échantillon	effectifs théoriques ≥ 5	<code>shapiro.test(x,...)</code>
χ^2 d'ajustement	1 échantillon		<code>chisq.test()</code>
Kolmogorov-Smirnov	1 échantillon		<code>ks.test(x,.)</code>
	2 échantillons		<code>ks.test(x,y)</code>
Tests de position :			
médiane	1 échantillon	$\min(n_1, n_2) \geq 10$ $\min(n_1, n_2) \leq 10$	<code>binom.test(x,)</code>
test du signe	2 échantillons		<code>fisher.test(x,y,)</code>
	2 éch. appariés		<code>binom.test(x,y,paired=T)</code>
Mann-Whitney	2 échantillons		<code>wilcox.test(x,y,exact=F)</code>
Mann-Whitney	2 échantillons		<code>wilcox.test(x,y)</code>
Wilcoxon	2 éch. appariés		<code>wilcox.test(x,y,paired=T)</code>