

Mikor	Mit
<ul style="list-style-type: none"> • Megfigyelt és várt gyakoriságokat hasonlítottunk össze. • Sok értéket tartalmaz (ha kevés, akkor is max csak hibaüzenet). 	<h2><u>χ²-próba (khi-négyzet)</u></h2> <p><u>lépései:</u></p> <ol style="list-style-type: none"> 1. c() -> c vektorba pakolni a méréseket 2. majd a c vektort egy táblába berakni (table) 3. chisq.test(megfigyelt gyak.->tábla, várt gyak.)
<ul style="list-style-type: none"> • Van egy elemi esemény és ezt akarjuk megismételni n-szer. 	<h2><u>Binomiális-próba</u></h2> <p><u>lépései:</u></p> <ol style="list-style-type: none"> 1. binom.test (kedvező kimenetek száma [például: x=15], összes ismételés [például: n=16], és valószínűsége [például: p=0,5])
<ul style="list-style-type: none"> • Kettőnél több független várható értékek tesztelése. 	<h2><u>Egyszempontos ANOVA</u></h2>
<ul style="list-style-type: none"> • Egy várható értékünk van. 	<h2><u>Egymintás t-próba</u></h2>
<ul style="list-style-type: none"> • Két várható értékünk van. 	<h2><u>Kétmintás t-próba, ha függetlenek a minták (férfi-nő) és páros t-próba, ha összefüggőek a minták (ugyan azok az emberek)</u></h2>
<ul style="list-style-type: none"> • Az i-edik pont x-koordinátája a standard normális eloszlás i/n kvantilise. • Az y-koordinátája a tapasztalati eloszlás i/n kvantilise, vagyis a rendezett minta i-edik eleme. 	<h2><u>QQ ábra</u></h2> <p><u>lépései:</u></p> <ol style="list-style-type: none"> 1. standardizálás -> std_változó = scale(változó) 2. qqnorm(std_változó) / qqline(std_változó)
<ul style="list-style-type: none"> • Normalitásvizsgálat: egy változó normális eloszlást követ-e. 	<h2><u>Egymintás Kolmogorov-Szmirnov-próba</u></h2> <p><u>lépései:</u></p> <ol style="list-style-type: none"> 1. ks.test(változó, "pnorm", mean(változó), sd(változó)) <p>mean: mintaátlag sd: szórás</p>
<ul style="list-style-type: none"> • Homogenitásvizsgálat: Két vagy több ismeretlen eloszlást, amelyek mindegyikéből 1-1, egymástól független minták vannak, hasonlítottunk össze egymással. 	<h2><u>Kétmintás Kolmogorov-Szmirnov-próba</u></h2> <p><u>lépései:</u></p> <ol style="list-style-type: none"> 1. ks.test(változó[<i>miszerint</i> == "étéke"], változó[<i>miszerint</i> == "értéke"]) <ul style="list-style-type: none"> • az érték lehet szám is, ekkor viszont nem kell " "
<ul style="list-style-type: none"> • Konfidencia intervallum. 	<ol style="list-style-type: none"> ➤ t.test(változó, conf.level = 0.95) -> ez double-be fogja visszaadni ➤ t.test(változó, conf.level = 0.95)\$conf.int -> ez int-be fogja visszaadni

• Geometriai eloszlás paraméterének becslése.	<ol style="list-style-type: none">1. Értékek felvétele egy vektorba [pl.: c(1:8)]2. Gyakoriság [pl.: c(158,...)]3. $m = \text{sum}(\text{értékek} * \text{gyakoriság}) / \text{sum}(\text{gyakoriság})$4. $p = 1/m$
---	--

Mit	H_0	H_A
Egyszempontos ANOVA	Mindegyik várható érték egyenlő.	Mindegyik várható érték különböző.
Egymintás t-próba	Amikor a várható érték egy konkrét szám. $\mu=c$ vagyis $\mu=155$	$\mu \neq 155$
Páros t-próba	$\mu_1 = \mu_2$ Összefüggő minták.	$\mu_1 \neq \mu_2$
Kétmintás t-próba	$\mu_1 = \mu_2$ Független minták.	$\mu_1 \neq \mu_2$
Egymintás Kolmogorov-Szmirnov-próba	$P(\chi < x) = F(x)$ A vizsgált változó eloszlásfüggvénye megegyezik-e egy hipotetikus változó eloszlásfüggvényével.	$P(\chi < x) \neq F(x)$ Nem egyezik meg.
Kétmintás Kolmogorov-Szmirnov-próba	$P(\chi < x) = P(Y < x)$ A két eloszlásfüggvény egyenlő.	$P(\chi < x) \neq P(Y < x)$ Nem egyenlő.
χ^2 -próba	Gyakoriságok valószínűsége egyenlők.	Nem egyenlők.
Binomiális-próba	Valaminek a valószínűsége. [pl.: $p = 1/6$]	$p \neq 1/6$

Mit akarunk?		Parancs
Beolvasni egy beépített adatbázisból.		1. library(datasets) 2. input = data.frame(adatbázisnév)
Hisztogram + beosztás.		hist(változó) vagy hist(változó, break = 20)
Kvantilis.		quantile(változó, probs = c(0.25, 0.5, 0.75))
Nem tartalmaz adatokat a cella.		változó2[is.na(változó)] vagy: sum(változó, na.rm = TRUE) /*na.rm = TRUE semmit tartalmazó cellák*/
Mintaátlag.		mean(változó)
Mintaátlag lekérdezése valami szerint.		tapply(mit, mi szerint, mintaátlag, egyéb(pl.: na.rm = TRUE))
Szórás.		sd(változó)
Standard hiba.		sd(változónév)/sqrt(length(változónév))
Egymintás t-próba.		t.test(válnév, mu = 155, conf.level = 0.95) mu: várható érték
Páros t-próba.		t.test(vált1, vált2, conf.level=0.95, paired = TRUE) paired: összefüggőek-e
1.F-próba.	Kétmintás t-próba 2 lépésben	1. var.test(vált1, vált2, conf.level = 0.9)
2.Kétmintás t-próba.		2. t.test(vált1, vált2, conf.level = 0.9, var.equal = TRUE) var.equal: Feltételezhetjük-e a varianciák azonosságát. (Az előzőből [1. var.testes eredményből])
1.Levene teszt.	Egyszempontos ANOVA 2 lépésben	1. leveneTest(vált ~ factor(vált2), center=mean) factor: ezt akkor kell, mikor egy szám és diszkrétte alakítjuk
2.Egyszempontos ANOVA		2. oneway.test(vált ~ factor(vált2), var.equal=TRUE) vagy: m1 = aov(vált ~ factor(vált2)) summary(m1) summary: leírás az m1-ről
Itt az előző kettőnél (Leveneteszt meg az Egyszempontos ANOVA) az első lépés szintén a varianciák azonosságának tesztelése, majd a második lépés a várható érték tesztelése.		
Ferdesség: Ha pozitív, akkor jobbra ferde eloszlás, szimmetrikus, ha negatív nyilván a fordítottja.		skewness
Lapultság: Ha pozitív, akkor ez az eloszlás a normális eloszlásnál csúcsosabb görbét kapunk, ha meg negatív, laposabb görbét kapunk.		kurtosis