Biostatisztika és informatika alapjai

4. előadás: Az orvostudományban előforduló nevezetes eloszlások

2017. október 5.

Veres Dániel

Ebben az előadásban az elméleti eloszlásokról lesz szó.
Először a statisztika alapvető problematikájáról: sokaság – minta – hiba beszélünk.
Feltesszük a kérdést, hogy hogyan következtethetünk egy minta tulajdonságaira a populációból. Ezt követően definiáljuk az elméleti eloszlásosokat és azok paramétereit.
Majd megtárgyaljuk, hogy milyen kérdésre melyik elméleti eloszlást használhatjuk fel: leírjuk az egyenletes, binomiális, geometriai, Poisson, Gauss, lognormál, exponenciális eloszlásokat példákkal. Ezt követően szót ejtünk a khí-négyzet és a t-eloszlásokról, majd végül az eloszlások transzformációira hozunk néhány példát.

Alapsokaság és minta

Alapsokaság
az elméletileg lehetséges összes mérést, vizsgálati eredményt tartalmazza

Minta
az alapsokaság egy részhalmaza, amit vizsgálunk

Ezen a dián ismételten felhívnám a figyelmet a statisztika alapvető problematikájára: az elméleti alapsokaság az elméletileg lehetséges összes vizsgálati eredménnyel (változók és kimenteleik) adott, azonban mi csak ennek egy részhalmazát – a véletlenszerűen vett mintát - vizsgáljuk adott körülmények között.
A következtető statisztika egyrészt a minta alapján következtet az alapsokaságra, másrészt a sokaságból a mintára. Bármelyik lehetőséget is vesszük, mindkét esetben fenn áll a bizonytalanság, amit a statisztikában valószínűségekkel jellemzünk.
Ebben az előadásban főleg azzal foglalkozunk, hogy a minta egyes paramétereire hogyan következtethetünk a sokaság alapján – de ez utóbbit is mintákból szoktuk becsülni.

Adott mintára vonatkozó felmerülő statisztikai kérdések az orvostudományban…

Az előző év alapján mekkora a valószínűsége, hogy a rendelkezésre álló 4 oltóanyag elegendő lesz, ha 25 embert várunk aznapra?

Hány szülés várható az esti ügyeletben, ha az éves statisztika 1000 szülést mutat éjfél és 8:00 között?

Az évfolyamból várhatóan hányan lesznek alkalmasak egy csípőprotézis elvégzésére (tömegük alapján)?

Mekkora a valószínűsége annak, hogy páciensünk 3.45 mmol/l-es K szintje még „egészséges”?

Az influenza vagy AIDS tesztünk pozitív – mekkora a valószínűsége, hogy valóban betegek vagyunk?

Milyen kérdések merülhetnek fel bennünk tehát? Ezen a dián soroltam fel néhány ilyen példát. Például az első kérdésnél a sokaságot az előző éves adatok jelentik, a vizsgált tulajdonság az oltásszükséglet, míg a minta az aznapra várt 25 páciens.

Milyen a megfelelő alapsokaság ?

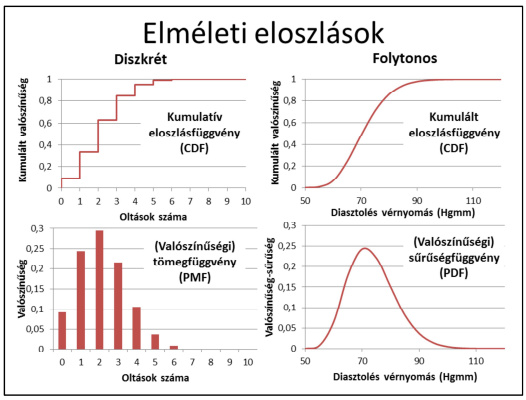
IGUÁNA
BÖLÉNY
STRUCC
PANDA
KATICABOGÁR

?

Hogyan etessük?

4

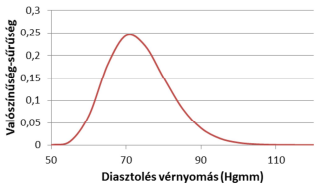
A minta adott tulajdonságára vonatkozó jó következtetéshez (vagy a sokaságra vonatkozó következtetéshez – ahogyan azt a későbbi előadásokon megfigyelhetjük) ismernünk kell a sokaságot – ahogyan a mintaállatunk hovatartozását (fáját – sokaságát) is tudnunk kell a megfelelő tápláláshoz (ami az adott vizsgált tulajdonságunk lehet).
Mit is jelent a sokaság ismerete?



Egy sokaság tökéletesen jellemezhető az eloszlásával (ahogyan a minta is, amint azt korábban már tanultuk). A sokaság eloszlását elméleti eloszlásnak nevezzük.
Az elméleti eloszlások valószínűségeket mutatnak gyakoriságok helyett – emlékezzünk csak a gyakoriságokra vonatkozó nagy számok törvényére, ahonnan tudjuk, hogy a relatív gyakoriság a valószínűséghez tart . Ha az elemszám végtelen, azaz a teljes sokaságot vizsgáljuk, akkor a relatív gyakoriság helyett valószínűségről beszélhetünk.
Az elméleti eloszlásokat 2 úton csoportosíthatjuk.

1. Ahogyan a változóknál, itt is beszélhetünk folytonos és diszkrét változó-eloszlásokról.
2. A minta leírásához hasonlóan az elméleti eloszlásoknál is beszélhetünk valószínűségekről és kumulált valószínűségekről – amelyek között az integrálás és deriválás matematikai művelet teremt kapcsolatot.

Az ábrán feltüntetem az egyes típusok statisztikai elnevezését is.

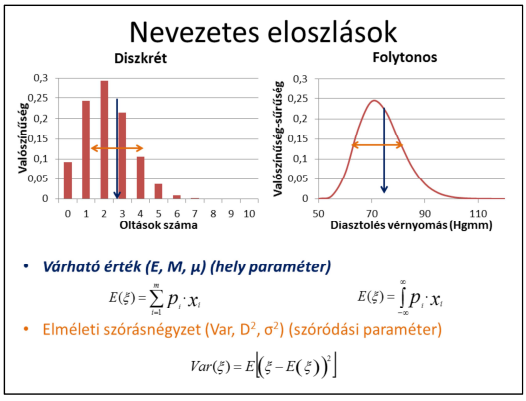


Ismerem az adott mennyiség valószínűségét minden esetre. (Nagyon-nagyon ritka – percentilis tábla.)

A valószínűséget ki tudom számítani *(vagy becsülni tudom)*: *nevezetes eloszlások paraméterei alapján.*

Milyen paramétereket, melyik eloszlást használjam?

A korábban említett kérdések megválaszolásában tehát segítenek minket az *elméleti eloszlások,* mivel megmutatják egy adott érték előfordulásának valószínűségét. Ritka esetben igen-igen nagy számú (végtelen) mérés alapján ismert egy adott változó valószínűségi eloszlása – ld. például percentilis táblázatok, görbék, de ez igen ritka.
Az esetek nagyobb részében azonban néhány jellemző érték és úgynevezett *nevezetes eloszlások* alapján tudjuk meghatározni bármely értékhez tartozó, így az adott kérdéses értékünk (pl. oltóanyagszám) valószínűségét is. Ebben az esetben a kérdés csak annyi, hogy *melyek ezek a néhány jellemző értékek* – hogyan tudom ezeket megadni, illetve becsülni –, illetve melyik kérdésemre *melyik eloszlást* kell használnom.



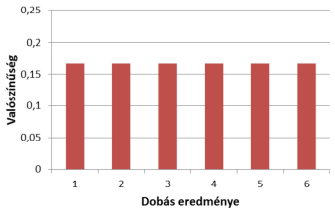
Vizsgáljuk meg először, hogy az elméleti eloszlásoknak melyek ezek a jellemzői, paraméterei.
Két alapvető jellemzőt használunk - a leíró statisztikában tanultakhoz hasonlóan: egy *közép értéket* és egy *szóródási paramétert.*
Kétféle eloszlást mutatok be az ábrán: egy folytonos (Hány Hgmm a diasztolés vérnyomása az embereknek) és egy diszkrét változó (hány oltás szükséges az előző év alapján) eloszlását.

A középérték jellemzésére a *várható értéket* (E, M vagy μ jelöléssel), a szóródás jellemzésére az *elméleti szórást* (Var, D^2 vagy σ^2 jelöléssel) használjuk. A várható érték az eloszlás „közepét" mutatja (ahogyan a minta esetén a módusz, a medián és az átlag – tehát ezek használatosak a becslésére). Az elméleti variancia az elméleti eloszlás „szélességét" mutatja (ahogyan a minta varianciája, illetve kvantilistávolságai). Ez a két jellemző egyértelműen leírja az általunk használt speciális eloszlásokat – azaz ezek ismeretében bármely értékhez tartozó valószínűség meghatározható.

Vizsgáljuk meg először a diszkrét esetben kapott várható érték definíciót. Vegyük észre, hogy p – a nagy számok törvényét használva – közelíthető a relatív gyakorisággal, amely pedig az abszolút gyakoriság osztva az elemszámmal. Tehát E felírható $E=(\Sigma(\text{abs.gyak.}_i*x_i))/n$, azaz az egyforma elemeket annyiszor adom össze, ahány van belőle (ez az abs.gyak._i*x_i), és ezt minden elemnél megteszem – tehát összességében minden elemet összeadok, majd az így kapott értéket osztom az elemszámmal. Vegyük észre, hogy ez nem más, mint az átlag definíciója végtelen elemszám esetében.
Folytonos változó esetében ugyanezt az összegzést végzem el, csak végtelenül kicsi

osztályszélességgel (ez az integrálás).
A következőkben megvizsgáljuk, hogy melyik speciális eloszlásnál hogyan adható meg a várható érték, illetve melyik eloszlás milyen kérdés megválaszolására alkalmas.
(Megjegyzés: a számunkra érdekes eloszlások várható értékei és szórásai szerepelnek a képlettárban.)

Egyenletes eloszlás



$$E(\xi)=\frac{1}{2}(a+b)$$

$$Var(\xi)=\frac{1}{12}(b-a)^2$$

$$Var(\xi)=\frac{(b-a+1)^2-1}{12}$$

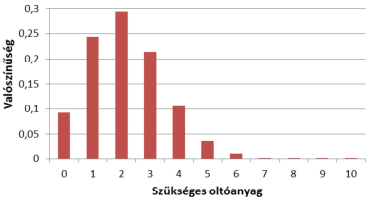
Ideális kocka eredményeinek eloszlása
Ideális munkaterhelés eloszlása a nap folyamán
Hőmérséklet eloszlása egy üres terem különböző pontjain

Először tekintsük meg az *egyenletes eloszlást*.
Egyenletes eloszlást mutat például az ideális kocka dobáseredményei, vagy az ideális napi munkaterhelés (minden másodpercben ugyanannyire kell dolgoznunk), illetve a hőmérséklet eloszlása egy üres terem különböző pontjain.

Az eloszlás alapján például megadhatom annak a valószínűségét, hogy 4-est dobok. A várható érték és a variancia képletében az *a* és a *b* a legkisebb, illetve legnagyobb érték. Ezek alapján tehát a 6 oldalú dobókockával való dobás várható értéke: 0,5*(1+6)=3,5.

Az előadás során bemutatott konkrét példát ikozaéder esetében a mellékelt excel fájl tartalmazza.

Binomiális (Bernoulli) eloszlás



$$E(\xi)=n\cdot p$$

$$Var(\xi)=n\cdot p\cdot(1-p)$$

$$P=\binom{n}{k}\cdot(p)^k\cdot(1-p)^{(n-k)}$$

Szükséges oltóanyagszám eloszlása
Általánosan: egy n-szer megismételt jelenség x-szer következik be
Ha p „kicsi” Poisson eloszláshoz „közelít”
Ha n „nagy” és p 0,5-höz tart, akkor normál eloszláshoz „közelít”

A *binomiális (Bernoulli) eloszlást* használunk általánosságban, ha egy dologot n-szer ismételve egy esemény k-szor következik be ez alatt.
A korábban említett példában az n – az ismétlések száma – a megvizsgálandó paciensek számát jelenti, a k pedig a raktáron lévő oltóanyagok száma.

Látható, hogy a várható érték, illetve a szórás (és így az adott értékhez tartozó egyenlet) meghatározásához szükséges tudnunk (vagy becsülnünk) a bekövetkezés valószínűségét – példánkban az oltóanyag szükségességének relatív gyakoriságát.
Megjegyzendő, hogyha az esemény bekövetkezésének valószínűsége (az ismétlésszámhoz képest) kicsi, akkor a binomiális eloszlás a Poisson eloszláshoz közelít.
Ha az ismétlések száma nagy és p értéke 0,5-höz tart, akkor pedig a normál eloszláshoz válik hasonlóvá.

Példaszámítás….

Influenzaszezont megelőzően a rendelőnkben az adott napra 4 oltóanyag áll rendelkezésre. Az előző években átlagosan 2989 páciensből 402 személyt kellett beoltanunk. Az előző év alapján mekkora a valószínűsége, hogy a rendelkezésre álló 4 oltóanyag elegendő lesz és el is fogy, ha 25 embert várunk aznapra?

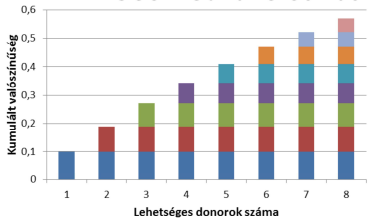
$$P=\binom{n}{k}\cdot(p)^k\cdot(1-p)^{(n-k)}=\binom{25}{4}\cdot\left(\frac{402}{2989}\right)^4\cdot\left(1-\frac{402}{2989}\right)^{(25-4)}\approx 0{,}2$$

Egyszerűbben lehetne? - Excel

A kérdés megválaszolásához a Bernoulli eloszlást (lásd később) használjuk fel (vegyük észre, hogy a feladat során a 25 emberből választunk ki 4-et jelenti az ismétlés nélküli kombinációt).

A példaszámolás eredményét a mellékelt excel fájl tartalmazza.

Geometriai eloszlás



$$E(\xi)=\frac{1}{p}$$

$$Var(\xi)=\sqrt{\frac{1-p}{p^2}}$$

$$P=p\cdot(1-p)^{(n-1)}$$

Független Bernoulli kísérletek egymásutánja
Hanyadikra találjuk meg a megfelelő donort? (Mekkora a valószínűsége annak, hogy az x. donorból megtaláljuk az első megfelelőt?)
Hanyadik szülésből lesz először fiú?

A *geometriai eloszlás* egy speciális Bernoulli eloszlás. Ezt az eloszlást kapjuk, ha egymástól független Bernoulli próbákat hajtunk végre egymás után (megtörténik az esemény, vagy nem történik meg).
Egy orvosi példa erre az eloszlásra: Mekkora a valószínűsége, hogy szükségünk van a nővér segítségére az első betegnél? Vagy csak a másodiknál kell hívnunk segítséget...
Hanyadikra találjuk meg a megfelelő donort? (Mekkora a valószínűsége annak, hogy az x. donorból megtaláljuk az első megfelelőt?)
Az ábrán azt tüntettem fel az első oszlopban, hogy mekkora a valószínűsége annak, hogy már az első donor megfelelő, a második oszlopban, hogy vagy az első vagy a második esetben lesz jó a donor... tehát itt egy kumulált valószínűséget tüntettem fel, amit úgy fogalmazhatunk meg, hogy annak a valószínűsége, hogy legkésőbb hanyadik donor lesz jó.

Ilyen eloszlást látunk a Szentpétervári paradoxonban is.

Péter és Pál

(pétervári paradoxon)

Pénzfeldobásos játék, eredménye szerint

- a kezdeti nyeremény 2 dukát és mindig duplázódik, ha írást dobunk
- ha fej, akkor vége a játéknak, írásnál folytatódik
- a nyeremények:
  - ha az 1. dobás fej: Pál ad 2 dukátot Péternek
  - ha az 1. írás, 2. dobás fej: Pál ad 4 dukátot Péternek
  - ha csak a 3. lesz fej: Pál ad 8 dukátot Péternek...

Mennyit fizessen Péter a játékélményért egy játékra átlagosan?
(Úgy hogy ne járjon se Pál, se Péter rosszul)

A valószínűségszámítás egyik kezdeti, a fejlődést megindító, meghatározó problémája az 1715-ben megjelent *pétervári paradoxon* volt. Ez egy pénzfeldobásos játék két elméleti játékos Péter és Pál között. A játék szabályai a következőek:

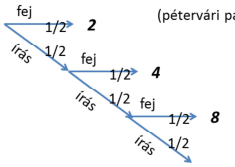
Pénzfeldobásos játék, eredménye szerint

- a kezdeti nyeremény 2 dukát és mindig *duplázódik*, ha írást dobunk
- ha fej, akkor vége a játéknak, írásnál folytatódik
- a nyeremények:
  - ha az *1. dobás fej*: Pál ad 2 dukátot Péternek
  - ha az 1. írás, *2. dobás fej*: Pál ad 4 dukátot Péternek
  - ha csak a *3. lesz fej*: Pál ad 8 dukátot Péternek...

Mennyit fizessen Péter a játékélményért egy játékra átlagosan? (Úgy hogy ne járjon se Pál, se Péter rosszul – mindketten a pénzüknél maradjanak)

Péter és Pál

(pétervári paradoxon)



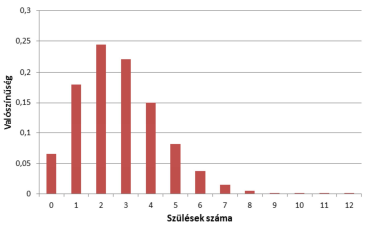
Az egy játékra jutó átlagos nyeremény matematikailag („elméletileg"): végtelen!

$$\frac{1}{2}\cdot 2+\frac{1}{2^2}\cdot 4+\ldots+\frac{1}{2^n}\cdot 2^n$$

Tapasztalat: Buffon 2084 játékból átlagosan 9,82 dukátot nyert egy játékra nézve.

Tehát mi is lesz egy játékban Pál „várható átlagos" adománya?
Az esetek felében az első dobás fej – ez *½*2* dukátot jelent (az egyenlet 1. tagja) Péter számára egy átlagos játékban. Ha az első írás (ez ugye az esetek fele), akkor ezen esetek felében fej lesz a következő dobás eredménye, tehát átlagosan egy játékra nézve ez *½*½* esetben történik meg, azaz *¼*4* dukátot jelent Péternek. Ezt a gondolatmenetet folytatva figyelembe véve az összes lehetséges esetet Péter ezeknek az összegét nyeri egy játékban átlagosan. Ha megvizsgáljuk ezt az összeget, akkor láthatjuk, hogy *elméletileg* Péter egy játékban *átlagosan 1+1+1+1.... azaz végtelen dukátot kap*!
Na de *a gyakorlatban* azt tapasztaljuk, hogy egyetlen játékban sem kap végtelent Péter, illetve az átlagos nyereménye *egy játékra soha nem végtelen*!
A paradoxon elnevezés tehát erre a tapasztalati-elméleti (nem végtelen-végtelen) ellentétre utal.
Példának okáért Buffon (egy híres matematikus) 2048 játékot játszott és 9,82 dukátos átlagnyereményt produkált (közel sem végtelent...) Egymillió játékból pedig 10,94 dukátos átlagot kaptak egy számítógépes modellezésben.
A tapasztalat azt mutatta, hogy bár Péter átlagos nyereménye egy játékra nézve sohasem végtelen, de a *játékok számának növelésével mindig egyre több, közelítve az elméleti értékhez*!

Poisson eloszlás



$$E(\xi)=\lambda$$

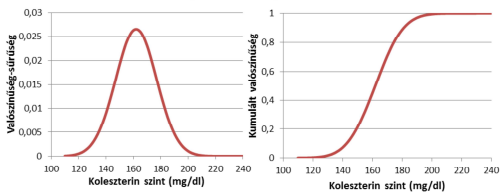
$$Var(\xi)=\lambda$$

$$P=\frac{\lambda^x}{x!}\cdot e^{-\lambda}$$

Születések száma az ügyeletben
Új diagnosztizált karcinómák száma
Hálóban lévő halak száma
Radioaktív preparátumban adott idő alatt elbomló atomok száma
Normál eloszláshoz „közelíthető"

Ez az eloszlás is a binomiális eloszlás egy különleges esete. A *Poisson eloszlásnak* van egy különlegessége: *várható értéke és szórása megegyezik*, valamint *egyetlen paraméterrel megadható*, így a számítások könnyebbek lehetnek.
Például ennek az eloszlásnak a segítségével megadhatjuk, hogy mekkora annak a valószínűsége, hogy 3 szülés történik az éjszakai ügyeletben. További példák az új diagnosztizált karcinómák számának, illetve a hálóban lévő halak számának eloszlása.
Általánosságban Poisson eloszlást követ a változónk, ha a kérdés egy adott elemszám egy adott intervallum alatt, illetve adott térfogatban, területen..., *ha az esemény bekövetkezésének valószínűsége kicsi*. (Példánkban a szülés valószínűsége kicsi – az emberek kis részével történik meg egy adott időben.)
A várható érték – a λ – az ismétlések számának és az adott esemény bekövetkeztének valószínűségének szorzatából adódik (n*p).
Megjegyzendő, hogy megfelelő körülmények esetében ez az eloszlás is normál eloszlással közelíthető.

Normál (Gauss) eloszlás I.



Koleszterinszin, vércukorszint....
Testmagasság, BMI
Diasztolés vérnyomás felnőtteknél
.....

$$E(\xi)=\mu$$

$$Var(\xi)=\sigma^2$$

$$P=\frac{1}{\sigma\cdot\sqrt{2\cdot\pi}}\cdot e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

A *normál (Gauss) eloszlásnak* a fentebb már említetteken kívül további különlegessége van: *ez az orvosi gyakorlatban a leggyakoribb eloszlás*.
Az ábrán az eloszlás sűrűségfüggvényét és kumulatív eloszlásgörbéjét egyaránt feltüntettem, ugyanis ez nagyon fontos eloszlás számunkra. Mint látható, az előzőekkel ellentétben a normál eloszlás egy szimmetrikus eloszlás.

A *normál (Gauss) eloszlásnak* a fentebb már említetteken kívül további különlegessége van: *ez az orvosi gyakorlatban a leggyakoribb eloszlás*.
Az ábrán az eloszlás sűrűségfüggvényét és kumulatív eloszlásgörbéjét egyaránt feltüntettem, ugyanis ez nagyon fontos eloszlás számunkra. Mint látható, az előzőekkel ellentétben a normál eloszlás egy szimmetrikus eloszlás.
Látható, hogy valóban sok a változó, amely normál eloszlást követ: a koleszterinszint, a vércukorszint, a legtöbb enzimszint, a testmagasság, a BMI, a disztolés vérnyomás.... De vajon miért van ez?

Normál eloszlás II.

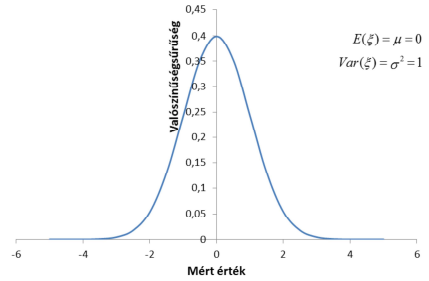
Centrális határeloszlás tétele (változókra): ha sok független valószínűségi változót összegzünk, akkor elég általános feltételek teljesülése esetén az összeg normális eloszlású valószínűségi változó lesz.

Centrális határeloszlás tétele (mintavételi átlagokra): ha egy adathalmazból *n* elemű mintákat veszünk, akkor elég általános feltételek teljesülése esetén a minták átlagai normál eloszlásúak lesznek, és az eloszlás varianciája az eredeti eloszlás varianciájának *n*-ed része lesz.

A normál eloszlás gyakoriságának okára a *centrális határeloszlás tétele* utal. Ez kimondja, hogy ha sok független valószínűségi változót összegzünk, akkor elég általános feltételek teljesülése esetén az összeg normális eloszlású valószínűségi változó lesz.
Az emberi test különböző jellemzői (mérhető értékei, változói) általában nagy sok más változó együtteséből alakulnak ki. Például az emberi testmagasság függ az apai és anyai génektől, a táplálkozástól, az életviteltől...

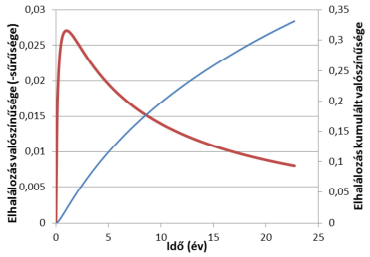
Erre vonatkozó példát tartalmaz a mellékelt excel fájl.
Másik megfogalmazása a centrális határeloszlás tételének a mintavétel során kapható átlagra vonatkozik: ha többször veszünk n elemű mintát egy sokaságból, mintából, akkor (ha n elég nagy), a minták átlagainak eloszlása normál eloszlás lesz, és ennek szórásnégyzete az eredeti szórásnégyzet n-ed része lesz.
Erre példát a következő előadásban láthatunk.

Standard normál eloszlás



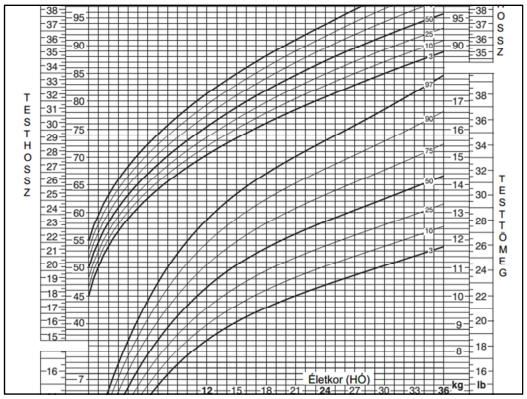
A standard normál eloszlás egy olyan normál eloszlás, aminek a várható értéke 0 és szórása 1.

Lognormál (Galton) eloszlás



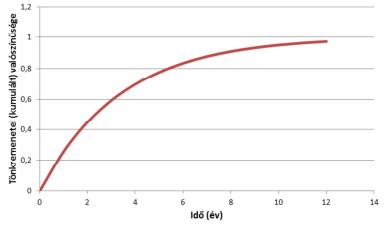
Testtömeg gyermekkorban
Túlélési idő

A l*ognormál* eloszlásnak is nagy jelentősége van az orvosi gyakorlatban. Ilyen eloszlást követ például a testtömeg, testmagasság gyermekkorban, illetve a túlélési idő rosszindulatú daganatoknál.
Általánosságban azt mondhatjuk, hogy akkor lesz az eloszlásunk lognormál, ha a változónk értékei kicsik, de nem lehetnek 0-nál kisebbek.
(Az eloszlást azért nevezzük lognormálnak, mert a változó értékeinek logaritmizálásával normális eloszláshoz jutunk.)



A percentilis görbékből is látható a lognormál eloszlás asszimetriája (látva, hogy pl. a 3-as és 10-es percentilis távolsága más, mint a 97-es és 90-es percentilisé.

Exponenciális eloszlás



$$E(\xi)=\frac{1}{\lambda}$$

$$Var(\xi)=\frac{1}{\lambda^2}$$

$$P=\lambda\cdot e^{-\lambda x}$$

Altatóberendezés működési ideje (az első hibáig).
Radioaktív bomlás során az egyes atomok élettartama.

Az *exponenciális eloszlás* igen gyakori a biofizikában és néhol van szerepe az orvosi gyakorlatban is. Ehhez két példát említenék: altatóberendezés működési ideje (az első hibáig eltelt idő), illetve radioaktív bomlás során az egyes atomok élettartama.

Változók transzformációi

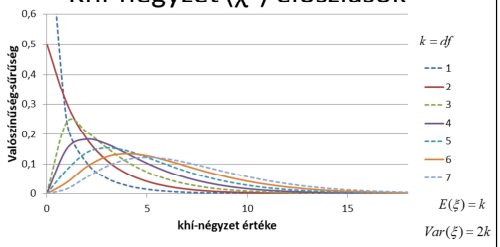
- Állandó hozzáadása
  $E(\eta)=E(\xi)+k \qquad Var(\eta)=Var(\xi)$
- Állandóval való szorzás
  $E(\eta)=E(\xi)*k \qquad Var(\eta)=Var(\xi)*k^2$
- Standardizálás
  Állandó hozzáadása, majd állandóval szorzás
  $$\eta-(\xi-E(\xi))*\frac{1}{\sqrt{Var(\xi)}}-\frac{(\xi-E(\xi))}{\sqrt{Var(\xi)}} \qquad \begin{matrix}E(\eta)=0\\Var(\eta)=1\end{matrix}$$
- Változók összeadása
  $E(\eta)=E(\xi)+E(\omega) \qquad Var(\eta)=Var(\xi)+Var(\omega) \leftarrow$ *függetlenségnél*
  Stabil eloszlás: ha az eloszlás ugyanaz marad
- Változók összeszorzása
  $E(\eta)=E(\xi)*E(\omega)$

21

A hatékony összehasonlításhoz gyakran transzformáljuk az adatokat. Itt a normál eloszlás egyes transzformáltjairól esik néhány szó.
A várható érték és a variancia változásai láthatók az ábrán állandóval való műveletek esetében.
Vegyük észre, hogy a standardizálás nem más, mint először a várható értékkel való csökkentés, majd a szórással való osztás.
Stabil az az eloszlás, ahol adott eloszlású változók összege ugyanolyan típusú eloszláshoz vezet.

Lásd a mellékelt excel fájlban levő példákat is.

Khí-négyzet (χ^2) eloszlások



$$E(\xi)=k$$

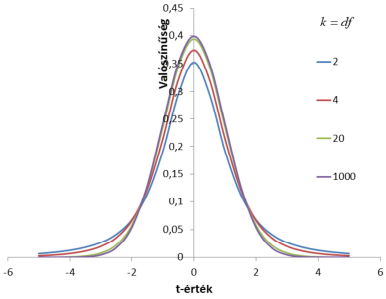
$$Var(\xi)=2k$$

Khí-négyzet érték: független, standard normál eloszlású változók négyzeteinek összege

22

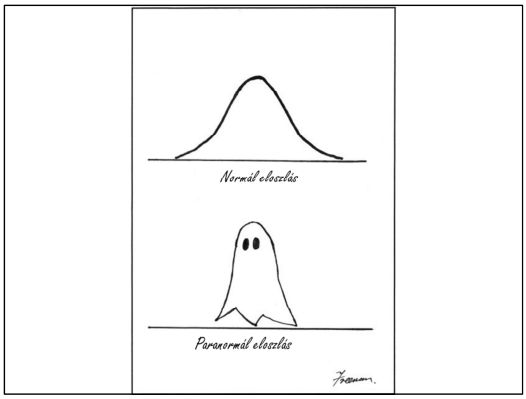
További 2 eloszlásról kell szót ejteni: a khí-négyzet és t-eloszlásokról.
Ezek az eloszlások a hipotézisvizsgálatoknál (lásd későbbi előadások) lényegesek. Mindkettő egy-egy eloszláscsaládot takar, ahol az eloszlások különbsége a szabadsági fokok számától (df) függ.
Khí-négyzet érték a független, standard normál eloszlású változók (k darab) négyzeteinek összege.
Az eloszlás várható értéke 2*k.

Student-féle t-eloszlás



23

A Student-féle t-eloszlás.
Ez egy szimmetrikus eloszláscsalád, ahol a várható érték mindig 0, szórása a szabadsági fokoktól függ. Végtelen szabadsági fok esetében megkapjuk a standard normál eloszlást.



Végezetül bemutatnék egy nem statisztikai eloszlást…

Ellenőrző kérdések #1

- Hogyan számítható egy folytonos eloszlás várható értéke?
- Hogyan számítható egy diszkrét eloszlás várható értéke?
- Melyik középérték egyezik meg a várható értékkel egy populáció esetében?
- Mivel becsülhető egy elméleti eloszlás várható értéke?
- Mivel becsülhető egy elméleti eloszlás szórása?
- Definiáld a z elméleti varianciát.
- Melyik két mutató határoz meg egyértelműen egy speciális eloszlást?
- Ábrázold az egyenletes eloszlás gyakorisságfüggvényét.
- Ábrázold a Poisson eloszlás gyakorisságfüggvényét.
- Ábrázold a Bernoulli eloszlás gyakorisságfüggvényét.
- Ábrázold a geometriai eloszlás gyakorisságfüggvényét.
- Ábrázold a normál eloszlás gyakorisságfüggvényét.
- Ábrázold a Gauss eloszlás kumulált eloszlásfüggvényét.
- Ábrázold az exponenciális eloszlás gyakorisságfüggvényét.
- Ábrázold a lognormál eloszlás gyakorisságfüggvényét.
- Írj 2 példát az egyenletes eloszlásra.
- Írj 2 példát binomiális eloszlásra.
- Írj 2 példát a geometriai eloszlásra.
- Írj 3 példát a normál eloszlásra.
- Írj 2 példát a lognormál eloszlásra.
- Írj 2 példát a Poisson eloszlásra.
- Írj 2 példát az exponenciális eloszlásra.
- Hogyan számítható az egyenletes eloszlás várható értéke?
- Hogyan számítható a binomiális eloszlás várható értéke?
- Hogyan számítható a lognormál eloszlás várható értéke?
- Hogyan számítható az exponenciális eloszlás várható értéke?
- Hogyan számítható a Poisson eloszlás várható értéke?
- Hogyan számítható normál eloszlás várható értéke?
- Miről szól a centrális határeloszlás tétele?
- Miért követ a legtöbb orvosi gyakorlatban használt változó normál eloszlást?
- Mik a standard normál eloszlás paramétereinek számszerű értéke?

Ellenőrző kérdések #2

- Add meg általánosan, hogy mikor kapunk általában binomiális eloszlást.
- Add meg általánosan, hogy mikor kapunk általában geometriai eloszlást.
- Add meg általánosan, hogy mikor kapunk általában Poisson eloszlást.
- Add meg általánosan, hogy mikor kapunk általában lognormál eloszlást.
- Milyen transzformációval kaphatunk a lognormál eloszlásból normál eloszlást?
- Hogyan kapunk khí-négyzet eloszlású változót?
- Definiáld a populációt és a mintát.
- Hogyan változik a változó eloszlásának várható értéke és a szórása konstans hozzáadására?
- Hogyan változik a változó eloszlásának várható értéke és a szórása konstanssal való szorzáskor?
- Hogyan lehet standardizálni egy eloszlást? Mit jelent ez?
- Hogyan változik a várható érték és a szórás független normál eloszlású változók összeadásakor?