A gyakorló feladatok számozása a bevezető órát követő órán, azaz a második órán indul.

Gyakorló feladatok megoldásai 1

1. A populációt a számunkra érdekes egységek (személyek, családok, iskolák stb.) alkotják, a mintát pedig azok, akiket valójában képesek vagyunk megvizsgálni.
2. Ha a populáció minden elemét vizsgáljuk.
3. Az ideális minta reprezentatív a populációra nézve. Kellően nagy (az elemszáma) és ha a mintavétel véletlen, független (random).
4. Random mintavétel esetén következtethetünk a populációra.
5. Kísérlet esetén levonhatunk oksági következtetéseket.
6. Mert kísérletet alapozhatnak meg; mert időnként nem akarunk oksági következtetést levonni; illetve nagyon ritkán mégiscsak képesek vagyunk megfigyelés esetén oksági következtetést levonni.
7. Statisztikai változó pl. az életkor és a 84 év egy adott értéke. Statisztikai változó a kedvenc állat is és annak egy lehetséges értéke a kutya.
8.
   a) A vizsgálni kívánt populáció a magyar általános iskolások összessége.
   b) 300, véletlenszerűen kiválasztott debreceni általános iskolás
   c) Nem, mert nem debreceni diákok nem kerülhettek be a mintába. Ezért a populációra vonatkozóan következtetéseket nem vonhatunk le.
   d) A debreceni általános iskolásokra vonatkozóan reprezentatív a minta.
   e) Két statisztikai változónk van: kreativitás, matematika tantárgy szeretete. Az előbbi bármely értéket felvehet 0 és 20 között, utóbbi pedig 1 és 20 között.
   f) Egy 300x2-es adatbázis lesz, minden sor egy személy, minden oszlop egy változó.
   g) Megfigyelésről van szó, ezért oksági következtetést nem vonhatunk le, ami itt egyébként se lenne logikus.

Gyakorló feladatok megoldásai 2

1. Sorba rendezhetőség, összegük értelmes, hányadosuk értelmes.
2. Pl. kedvenc szín, csillagjegy (nincs legkorábbi), város, ahonnan származik a személy
3. Az értékei sorba rendezhetőek, de az értékek különbségeik nem értelmesek.
4. Pl. intelligencia teszt érték, hőmérséklet
5. Az értékeinek hányadosai értelmesek.
6. Nem.
7. Igen.
9. Igen.
10. Igen.
11. Jól elkülönülő értékeik vannak.
12. A változó értékei mennyiséget (számszerűsíthető minőséget) fejeznek ki.
13. Nem diszkrét, hanem folytonos, és kvantitatív.
14. Diszkrét, és nem kvantitatív, hanem kvalitatív.
15. Ilyen a minimum, maximum, átlag, szórás, terjedelem stb.
16. A relatív gyakoriság a gyakoriság osztva a minta elemszámával. A gyakoriság azt fejezi ki, hogy hányszor fordul elő egy adott érték, míg a relatív gyakoriság azt fejezi ki, hogy milyen arányban fordul elő az adott érték a mintában.
17. A legtöbben a fehérek vannak és a legkevesebben az ázsaiak.

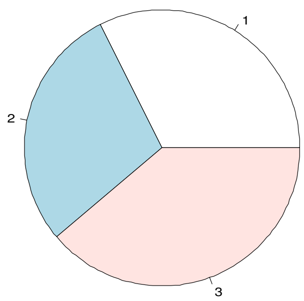
18. Középiskolai végzettségűek a leggyakoribbak (kb. 0.4, azaz 40%), PhD-vel rendelkezők a legritkábbak (kb. 0.025, azaz 2,5%).
19. A minta legnagyobb része 0 és 20 szál között szív cigarettát naponta, és kb. 260-an vannak.
20. A minta legkisebb része 60 és 80 szál között szív cigarettát naponta, és 10-nél kevesebben vannak.
21. Tegyük fel C: meghajtóra mentetted az adatbázist!

a=read.table("C:\\aPicasso.txt", sep=",", header=T)
a=as.matrix(a)

a. table(a[,2])
(Eredmény, ami nem volt kérdés:
1 2 3
100 89 120)

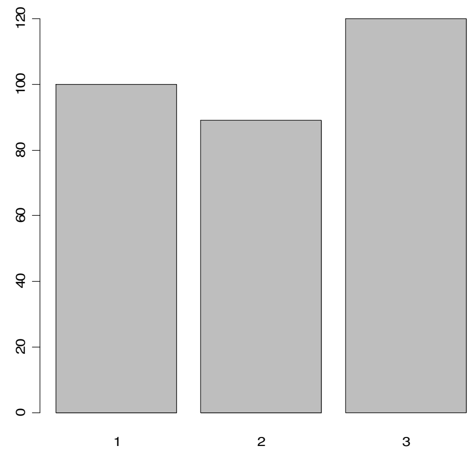
b prop.table(table(a[,2]))
(Eredmény, ami nem volt kérdés:
1 2 3
0.3236246 0.2880259 0.3883495)

c A legegyszerűbben, átalakítások, csinosítás nélkül:
pie(table(a[,2]))
(Eredmény, ami nem volt kérdés:



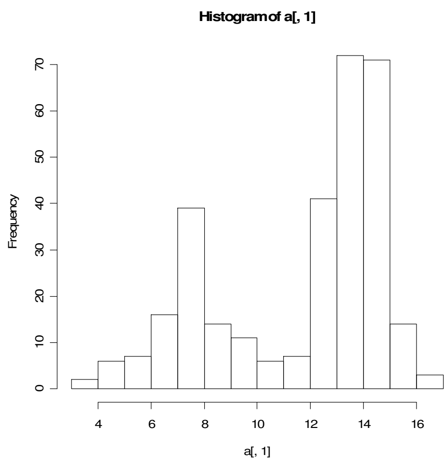
)

d A legegyszerűbben:
barplot(table(a[,2]))
(Eredmény, ami nem volt kérdés:



)

e A legegyszerűbben:
hist(a[,1])



Gyakorló feladatok 3 megoldásai

1. Igen
2. Nem
3. Nem
4. Igen
5. Bimodális

6. Módusz:6, medián:6 és az átlag: kb. 6.36.
   Terjedelem:4, az interkvartilis félterjedelem:0.5, a variancia: kb 1.478 és a szórás: kb. 1.22.

7. Módusz:kb. (170.55-164.89-)/2=2.83, a medián:170.15 és az átlag:170.12.
   Terjedelem: 127.74, az interkvartilis félterjedelem:6.95, a variancia:105.96 és a szórás:10.29.

Gyakorló feladatok megoldásai 4

1. P=14!
2. $C_5^{90} = \binom{90}{5} = \frac{90x89x88x87x86}{5x4x3x2x1} = 43949268$,

   illetve $C_2^{85} C_3^5 = \binom{85}{2}\binom{5}{3} = \frac{85x84x5x4x3}{2x1x3x2x1} = 35700$.
3. $V_2^{5,(ism)}$ =2^5 =32
4. $P(ism) = \frac{6!}{3!2!} = \frac{6x5x4x3x2x1}{3x2x1x2x1} = 60$.
5. $V_3^{10,(ism)} = 3^{10} = 59049$.
6. $C_7^{5,(ism)} = \binom{5+7-1}{7} = \frac{11x10x9x8x7x6x5}{7x6x5x4x3x2x1} = 330$.
7. $V_{30}^{40} = \frac{40!}{(40-30)!} = \frac{40!}{10!}$.
8. $C_2^3 C_2^4 = \binom{3}{2}\binom{4}{2} = \frac{3x2x4x2}{2x1x2x1} = 12$.
9. $V_4^{31} = 31 \cdot 30 \cdot 29 \cdot 28 = 755160$.
10. P=5!=5x4x3x2x1=120.
11. $C_3^{15,(ism)} = \binom{15+3-1}{3} = \frac{17x16x15}{3x2x1} = 680$.
12. $P(ism) = \frac{9!}{5!4!} = \frac{9x8x7x6x5x4x3x2x1}{5x4x3x2x1x4x3x2x1} = 1260$

Gyakorló feladatok 5

1.
   - a Elemi esemény
   - b Esemény
   - c Kísérlet
   - d Esemény, biztos esemény
   - e Kísérlet
   - f Esemény
   - g Elemi esemény
   - h Esemény, lehetetlen esemény
   - i Elemi esemény
   - j Kísérlet
   - k Elemi esemény
   - l Esemény, biztos esemény

2. 1/6
3. 3/7
4. 6/11
5. S={F,I}, p(F)=0.5
6.
   - a {1,2,3}, p=0.5
   - b {1,3,5}, p=0.5

c {3,4,5,6}, p=2/3
d {1,4,5,6}, p=2/3

7. 4/52. 8/52. P(treff dámaU király)= P(treff dáma)+ P(király)= 1/52+4/52=5/52
8. $P(\overline{A}) = 1 - p(A) = 1 - 0.35 = 0.65$,
   P(AUB)= P(A)+ P(B)- P(A∩B)= 0.35+0.51-0.17=0.69,
   P(S)=1, és P(Ø)=0
9. Független események esetén P(A∩B)= P(A)*P(B)=7/10*7/10=0.49
10. Független események esetén P(A∩B∩C)= P(A)*P(B)*P(C)=0.7*0.9*0.8=0.504
11. 6/15,10/15, és független események esetében P(AUB)= P(A)+ P(B)=6/15+5/15=11/15
12. P(AUB)= P(A)+ P(B)- P(A∩B)=4/6+3/6-2/6=5/6
13. 0.39/0.48=0.8125
14. 0.5/0.7=0.71
15. 0.3x0.8=0.24

Gyakorló feladatok 6

1. p(borravaló) = p(borravaló | kolumbiai)*p(kolumbiai) + p(borravaló | bécsi)*p(bécsi) + p(borravaló | guatemalai)*p(guatemalai)=
   0.5*0.2+0.2*0.7+0.3*0.1=0.27
   p(kolumbiai | borravaló) =
   [p(borravaló | kolumbiai)*p(kolumbiai)]/ [p(borravaló | kolumbiai)*p(kolumbiai) + p(borravaló | bécsi)*p(bécsi) + p(borravaló | guatemalai)*p(guatemalai)]=
   0.5*0.2/(0.5*0.2+0.2*0.7+0.3*0.1)=0.37
2. p(selejt) = p(selejt | 1. gép)*p(1. gép) + p(selejt | 2. gép)*p(2. gép) + p(selejt | 3. gép)*p(3. gép)=
   0.05*0.4+0.05*0.3+0.1*0.3=0.065
   p(1. gép | selejt) =
   [p(selejt | 1. gép)*p(1. gép)/ [p(selejt | 1. gép)*p(1. gép) + p(selejt | 2. gép)*p(2. gép) + p(selejt | 3. gép)*p(3. gép)]=
   0.05*0.4/(0.05*0.4+0.05*0.3+0.1*0.3)=0.31
3. p(1. doboz)=1/6=0.17 p(piros | 1.doboz)=3/4=0.75
   p(2. doboz)=2/6=0.33 p(piros | 2. doboz)=1/4=0.25
   p(3. doboz)=3/6=0.5 p(piros | 3. doboz)=2/4=0.5
   p(1.doboz | piros)=p(piros | 1.doboz)*p(1. doboz)/[p(piros | 1.doboz)*p(1. doboz)+ p(piros | 2.doboz)*p(2. doboz)+ p(piros | 3.doboz)*p(3. doboz)]=0.75*0.17/[0.75*0.17+0.25*0.33+0.5*0.5]=0.27

Gyakorló feladatok 7

Az itt szereplő megoldási utak szubjektív döntés eredményei, a válaszok viszont más megoldási mód esetén is nagyjából (a kerekítések miatt) azonosak.

1. Egyesével le lehet olvasni a megfelelő valószínűségeket a táblázatból, vagy lehet kérni a valószínűségeket a dbinom(k, N, p) paranccsal, ahol k az egyesek száma, N a kísérletek száma, p az egyessel kódolt érték valószínűsége. Ennek megfelelően a dbinom(0,10,0.5), dbinom(1,10,0.5), dbinom(2,10,0.5) stb. parancsokkal mindig egyel növelve az első számértéket megkapjuk a kívánt eloszlást. De akár egy lépésben is eljuthatunk az eredményhez a következő paranccsal:

   A megfelelő eloszlás:

Fejek száma	0	1	2	3	4	5	6	7	8	9	10
Valószínűség	0.001	0.01	0.04	0.12	0.21	0.25	0.21	0.12	0.04	0.01	0.001

2. dbinom(1, 3, 1/6), azaz 0.35
3. pbinom(2, 6, 1/6), azaz 0.94
4. dbinom(3, 8, 2/6), azaz 0.27
5. a 0, mert nincs 13-nál több pikk
   b pbinom(9,17,0.75)-pbinom(8,17, 0.75), azaz 0.028
   c dbinom(4,17,1/13), azaz 0,029
   d 0, mert nincs négynél több
   e pbinom(2,17,1/13), azaz 0,862
   f 1-pbinom(1,17,5/13), azaz 0,997
6.
   a pbinom(29,50,0.6) adja annak a valószínűségét, hogy 29 vagy annál kevesebb lány legyen, 1-pbinom(29,50,0.6) pedig a keresett valószínűséget. Értéke kb. 0.56.
   b pbinom(17,50,0.4), a valószínűségi érték kb. 0,24
   c 1-pbinom(29, 50, 0.6), azaz 0.56
   d pbinom(30, 50, 0.4), azaz 0,999
   e pbinom(35,50,0.6)-pbinom(25, 50, 0.6), azaz 0,848
   f dbinom(12, 50, 0.6), azaz $2x10^{-7}$
   g 50*0.4, azaz 20
7.
   a dbinom(20,20,0.75), azaz 0.003
   b pbinom(2,20,0.75), azaz $1.6x10^{-9}$
   c 1-pbinom(3, 20, 0.75), azaz kb. 1

8. a dbinom(1,25,0.3), azaz 0,001
   b dbinom(3,25,0.3), azaz 0,0243
   c dbinom(25,25,0.3), azaz $8,47x10^{-14}$
   d 1-pbinom(1, 25, 0.3), azaz 0,998
   e pbinom(15, 25, 0.3), azaz 0,9995
   f 25x0,3, azaz 7.5
9.
   a 1-pbinom(2,5,0.3), azaz. 0,472.
   b dbinom(5,5,0.7), azaz 0,168
   c dbinom(5,5,0.3), azaz 0,002
   d pbinom(3,5,0.3), azaz 0,969

Gyakorló feladatok 8

A Theodore Horvath Basic statistics for behavioral sciences című könyvében található feladatok nyomán.

1. A normál eloszlás területének mekkora hányada esik a
   a 1-pnorm(1.3,0,1), mert 1.3 az adott érték, 0 az átlag és 1 a szórás. És az 1.3-as értéktől nagyobb értékek valószínűségét keressük, ezért kell 1-ből kivonni annak valószínűségét, hogy 1.3 vagy ennél kisebb. 0.097 a terület aránya.

- b pnorm(-1.2,0,1), normál eloszlásnál bármely adott érték valószínűsége gyakorlatilag nulla, ezért nem baj, hogy a -1.2 benne marad a tartományban. Kb. 0.115.
- c 1-pnorm(2.1,0,1), azaz 0.018
- d 1-pnorm(-2.7,0,1), azaz 0.997
- e pnorm(1,0,1), azaz 0.84
- f pnorm(-1.96,0,1), azaz 0.025
- g A z=2-es érték az átlagtól jobbra esik. Az átlagtól jobbra eső terület aránya 0.5, ezért pnorm(2,0,1)-0.5, azaz 0.477
- h A z=-0.5-s érték az átlagtól balra esik, az átlagtól jobbra eső terület nagysága 0.5, így 1-pnorm(-0.5,0,1)-0.5, azaz 0.19
- i Fontos, hogy a nagyobb számértékhez tartozó valószínűségi értékből vonjuk ki a kisebbhez tartozót! pnorm(0.3,0,1)-pnorm(-0.3,0,1), azaz 0.24
- j pnorm(2.2,0,1)-pnorm(-2.2,0,1), azaz 0.97
- k pnorm(-0.35,0,1)-pnorm(-1.1,0,1), azaz 0.23
- l pnorm(1.36,0,1)-pnorm(0.36,0,1), azaz 0.27
- m pnorm(1.8,0,1)-pnorm(1.3,0,1), azaz 0.061
- n pnorm(1.96,0,1)-pnorm(-1.96,0,1), azaz 0.95

2. Melyik az a (két) z-érték, amely a görbe alatti terület
   - a Az átlagtól balra és jobbra elhelyezkedő értékek aránya 0.5. Az az érték, amely maga és az átlag között 25%-át tartalmazza az értékeknek, az az első vagy a harmadik kvartilist határozza meg. qnorm(0.25,0,1) és ugyanennek a számértéknek az abszolút értéke a szimmetria miatt. Vagy számolható a qnorm(0.75,0,1) paranccsal is. A két megfelelő érték tehát:-0.674, 0.674
   - b Az a legkisebb érték, mely maga és az átlag közt az adatok 45%-át tartalmazza, attól balra eső érték az adatok 5%-t tartalmazza. Ezért qnorm(0.05,0,1) , és a két megfelelő érték:-1.645, 1.645.
   - c qnorm(0.47,0,1) , és akét megfelelő érték:-0.075, 0.075.
   - d qnorm(0.65,0,1), azaz 0.385
   - e qnorm(0.15,0,1), azaz -1.036
   - f Azaz az átlag és az értékek közt 40%-a esik az adatoknak, így qnorm(0.1,0,1), és a két megfelelő érték: -1.28,1.28.
   - g qnorm(0.425,0,1), azaz a két megfelelő érték -0.19 és 0.19.
3. Az aránya azoknak, akik Dávidnál kevesebb pontot szereztek, de még átmentek: pnorm(56,60,5)-pnorm(50, 60,5), azaz 0.189. 40-en voltak a csoportban, azaz 0.189x40, azaz kb. 8-an voltak ilyenek.
4. A kérdés tulajdonképpen az, hogy mely értékek határolják ezen eloszlás középső 4%-t. qnorm(0.48, 162,6) és qnorm(0.52, 162,6), azaz a megfelelő értékek: 161.7 és 162.3.
5. qnorm(0.98, 100,15), azaz 130.8
6.
   a) Azaz az 55-ös értéknél magasabb érték valószínűségére vagyunk kíváncsiak a férfiaknál: 1-pnorm(55,65,15), azaz 0.75
   b) Férfiaknál 1-pnorm(75,65,15), azaz 0.252, nőknél 1-pnorm(75,55,10), azaz 0.023. Ugyanannyi férfi van, mint nő, ezért egyszerűen vehetjük az átlagot: 0.138, ennyi a keresett valószínűség.
   c) A nők 95%-a alacsonyabb értéket határoz meg, mint qnorm(0.95,55,10), azaz 71.45. A keresett valószínűség 1-pnorm(71.45, 65, 15), azaz 0.33
7.

a. Két egymástól független valószínűség, ezek együttes előfordulását keressük. A valószínűségek: pnorm (50,75,15), azaz 0.048, az együttes előfordulás ennek négyzete, azaz: 0.002
b. Egyenként pnorm(90,75,15)-pnorm(70,75,15), azaz 0.47, együttes előfordulás: 0.22
c. Egyenként pnorm(80,75,15), azaz 0.63, együttes előfordulás: 0.398

Gyakorló feladatok 9

1. μ=30, σ=8 (N(30,8)), N=16.
   $$SE = \frac{\sigma}{\sqrt{N}} = \frac{8}{\sqrt{16}} = \frac{8}{4} = 2$$
2. N=180, $\bar{x} = 12, s = 2$
   $$S\hat{E} = \frac{s}{\sqrt{N}} = \frac{2}{\sqrt{225}} = \frac{2}{15} = 0.13$$
3. a=c(7,11,6,8,10,9,5,2,13,13,7,9,9,5,11,10)
   sd(a)=3
   length(a)=16
   $$S\hat{E} = \frac{s}{\sqrt{N}} = \frac{3}{\sqrt{16}} = \frac{3}{4} = 0.75$$
4. N=200, n=20, N(23,4), SE=1.5
   a. (0.5-pnorm(21.5,23,1.5))*2=0.68
   b. qnorm(0.025,23,1.5)=20.06, qnorm(0.975,23,1.5)=25.94

Gyakorló feladatok 10

1. Például
   H0: A szorongási értékek terápia előtt és terápia után nem különböznek.
   H1: A szorongási értékek terápia után alacsonyabbak.
   Vagy ugyanez másképp
   H0: A terápiának nincs hatása.
   H1: A terápia hatásos.
   De kétoldali ellenhipotézist is lehet alkalmazni:
   H0: A szorongási értékek terápia előtt és terápia után nem különböznek.
   H1: A szorongási értékek terápia előtt és terápia után különböznek.

2. Például
   H0: A két terápia szorongáscsökkentő hatása nem tér el egymástól.
   H1: A két terápia szorongáscsökkentő hatása eltér.
   Vagy
   H0: Az új és régi terápia szorongáscsökkentő hatása megegyezik.
   H1: Az új terápia jobban csökkenti a szorongást, mint a régi.

3. H0: A nők és a férfiak fizetése átlagosan megegyezik a vizsgált vállalatnál.
   H1: A vizsgált vállalatnál a nők fizetése átlagosan alacsonyabb, mint a férfiaké.

4. H0: A nők és a férfiak fizetése átlagosan megegyezik a vizsgált vállalatnál.
   H1: A nők és a férfiak átlagfizetése átlagosan különbözik a vizsgált vállalatnál.

5. Igy nevezzük azt a döntést, ha a valóság az, hogy H0 hamis és H0-t elutasítjuk.
6. Igy nevezzük azt a döntést, ha a valóság az, hogy H0 igaz és H0-t megtartjuk.
7. Igy nevezzük azt a döntést, ha a valóság az, hogy H0 igaz és H0-t elutasítjuk.
8. Igy nevezzük azt a döntést, ha a valóság az, hogy H0 hamis és H0-t megtartjuk.

9. H0-t elvetjük.
10. H0-t elvetjük.
11. H0-t megtartjuk.
12. H0-t elvetjük.
13. 0.05, azaz 5%.
14. 0.99, azaz 99%
15. Ha a szignifikancia szint 0.5 lenne.

16.

$$u = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{N}}} = \frac{17000 - 15000}{\frac{5000}{\sqrt{25}}} = \frac{2000}{\frac{5000}{5}} = \frac{2000}{1000} = 2$$

A 2-es z értékhez tartozó kumulatív valószínűségi érték kb. 0.9772, azaz kétoldali ellenhipotézis mellet a valószínűség (1-0.9772)*2=0.0228*2=0.0556, azaz az általában alkalmazott 0.05-ös szignifikancia szint mellett az eredmény nem szignifikáns. Mindezek alapján nincs szignifikáns különbség a bányászok és a populáció fizetése között.