2.3. VERJETNOST

1. Zbirka 50 kamnin vsebuje 3 primerke granita. Če naključno izberem kos, kolikšna je verjetnost, da bo to granit?
2. Kakšna je verjetnost, da bo iz neke plasti naključno izbran amonit presegal 6 cm?
3. Vemo, da je nivo podtalnice v globini med 0 in 100 m, vendar ni znano kje. Kakšna je verjetnost, da bo v prvih 20 m pod površjem?
4. Podjetje mora dokončati predor v dveh leti ali pa bo moralo plačati globo. Kakšna je verjetnost, da bodo plačali globo.

2.3.1. Verjetnost - razlage

- Klasična razlaga

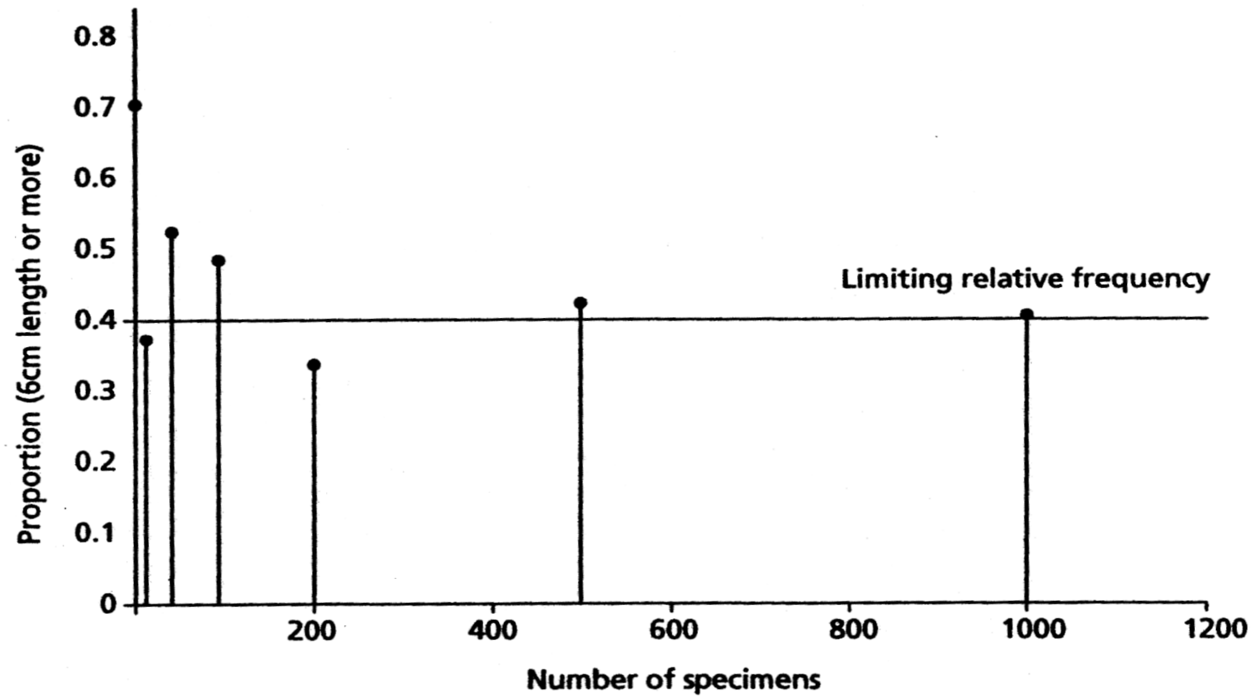
  Kadar ima poizkus n možnih, medsebojno izključujočih se izidov, med katerimi n_A določa dogodek A, tedaj je verjetnost (Pr), da se bo ta zgodil:

$$\Pr(A) = \frac{n_A}{n}$$

- Primer 1: 3/50 = 0,06 = 6%

2.3.1. Verjetnost - razlage

- Razlaga relativne pogostnosti
  Verjetnost dogodka E je vrednost, h kateri teži (limitira) njegova relativna pogostnost v daljšem zaporedju neodvisnih, naključnih poizkusov.
- Primer 2

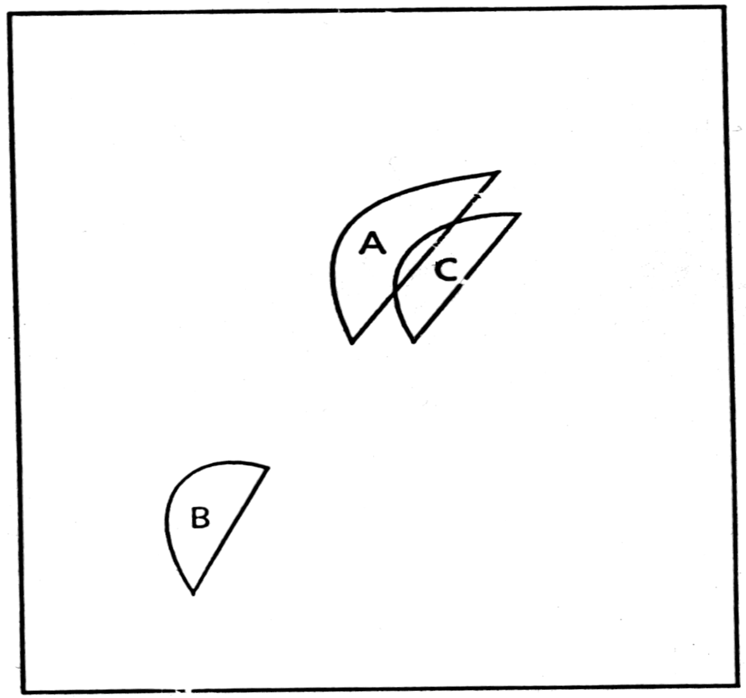


2.3.1. Verjetnost - razlage

- Geometrična razlaga

  Kadar je velikost celotnega razpoložljivega prostora *V*, je verjetnost, da leži naključno izbrana točka v območju, ki je veliko *v*

$$\Pr = \frac{v}{V}$$



- Primer 3:

  20/100 = 0,2 = 20%

2.3.1. Verjetnost - razlage

- Subjektivna verjetnost

  Verjetnost razlagamo kot stopnjo zaupanja. V katerihkoli okoliščinah imamo lastno predstavo o verjetnosti nekega dogodka; odvisna je od naših izkušenj (znanja) in načina, kako ga uporabimo $\Rightarrow$ Bayesian-ske statistike.

- Primer 4: Predor Šentvid ☹

2.3.2. Razpon vzorca

- Ne glede na to, katero interpretacijo uporabimo, moramo dognati vse možne izide poizkusa – razpon vzorca; njegova narava določa način izračuna verjetnosti dogodka.

POIZKUS	VRSTA IZIDA	RAZPON VZORCA
Izberi en kos izmed 50	Kateri kos?	Spisek 50 kosov
Izberi 5 kosov izmed 50	Število kosov granita?	0, 1, 2, 3
Opazuj poplave Mure v 25 letih	Število poplav?	0, 1, 2,...
Določi % SiO_2 v kamnini	% SiO_2?	Vse vrednosti med 0 in 100

2.3.3. Lastnosti verjetnosti
2.3.3.1. Merilna lestvica

- Razpon razmerij za opisovanje klasične in geometrijske verjetnosti ter verjetnosti relativne pogostnosti sega od 0 do 1. Dogodek E ima verjetnost:

$$0 \leq Pr(E) \leq 1$$

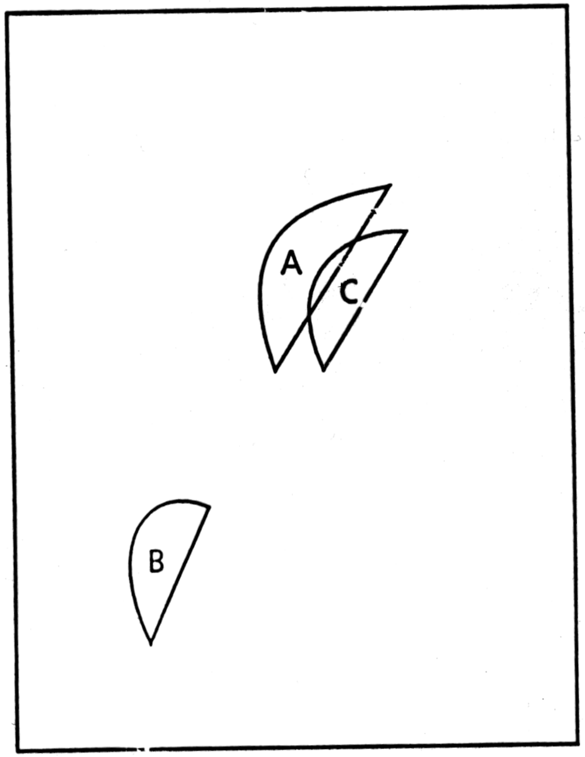
- Dogodek, ki ni možen ima verjetnost 0, tisti, ki se bo zagotovo zgodil 1.
- Gotovo je, da se bo pojavil eden izmed izidov, ki jih zajema razpon vzorca: celotna verjetnost vseh točk v razponu vzorca je 1.

2.3.3.2. Zakon seštevanja med seboj izključujočih se dogodkov

- Če so E_1, E_2, ..., E_k medsebojno izključujoči se dogodki, je verjetnost, da se dogodi en izmed njih, enaka seštevku njihovih posameznih verjetnosti.

$$Pr(E_1 \text{ ali } E_2 \text{ ali} \ldots \text{ ali } E_k) = Pr(E_1)+Pr(E_2)+\ldots+ Pr(E_k)$$

2.3.3.2. Zakon seštevanja med seboj izključujočih se dogodkov



- Primer 1a: Kakšna je verjetnost, da vrtina naključno zadane vodno telo A ali B?

$$\Pr(vrtina\_v\_A\_ali\_B) = \frac{površina\_A + površina\_B}{celotna\_površina}$$

- Primer 1b: Kakšna je verjetnost, da vrtina naključno zadane vodno telo A ali B?

$$\Pr(vrtina\_v\_A\_ali\_C) = \frac{površina\_A + površina\_C - površina\_prekrivanja\_AC}{celotna\_površina}$$

2.3.3.2. Zakon seštevanja med seboj izključujočih se dogodkov

Primer 2: Kakšna je verjetnost, da izhaja primerek vrste A iz nahajališča Y?

(št. primerkov vrste A) + (št. primerkov iz nahajališča Y) –
- (št. primerkov, ki ustreza obema pogojema)/100

	A	B	C	Σ
X	15	23	11	49
Y	24	14	13	51
Σ	39	37	24	100

(39+51-24)/100 = 0,66 = 66%

2.3.3.3. Komplementarni dogodki

- Verjetnost, da se dogodek E ne zgodi je enka *1 – Pr(E).*
- "dogodek E se ne zgodi" je komplementaren pojavu dogodka E. Označimo ga z $\bar{E}$ ali E_c.
- Pravilo verjetnosti medsebojno izključujočih se dogodkov:

$$Pr(\bar{E}) = 1 - Pr(E)$$

2.3.4. Pogojna verjetnost

- Pogojna verjetnost dogodka A, kadar veljajo okoliščine B je:

$$\mathrm{Pr(A|B)} = \frac{\mathrm{Pr}(A in B)}{\mathrm{Pr}(B)}$$

2.3.4. Pogojna verjetnost

Vsebnost Cu	A	B	C	Σ
Nizka	30	19	8	57
Srednja	10	20	17	47
Visoka	15	18	13	46
Σ	55	57	38	150

Primer:

1. Kakšna je verjetnost, da ima naključno izbran vzorec visoko vsebnost Cu?

   Pr(V) = 46/150 = 0,31

2. Kakšna je verjetnost, da ima naključno izbran vzorec visoko vsebnost Cu, če vemo, da je iz lokacije A?

   $$\Pr(V|A) = \frac{\frac{15}{150}}{\frac{55}{150}} = \frac{15}{55} = 0{,}27$$

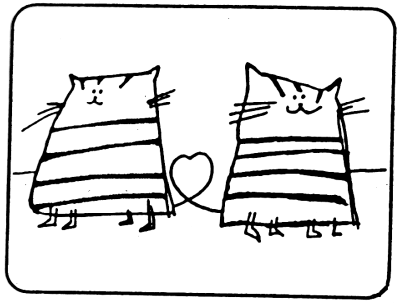
2.3.4.1. Zakon množenja povezanih dogodkov

- Pojaviti se morata tako dogodek A, kot tudi neke določene okoliščine B.

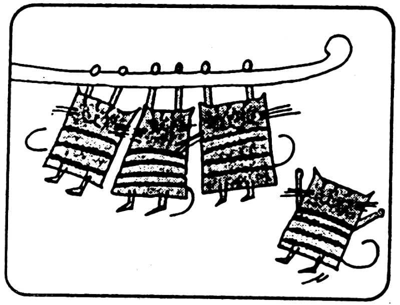
$$Pr(A \text{ in } B) = Pr(A|B) \cdot Pr(B)$$ oz.

$$Pr(A \text{ in } B) = Pr(B|A) \cdot Pr(A)$$

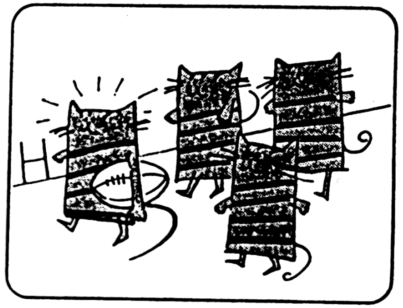
Štiri mucke



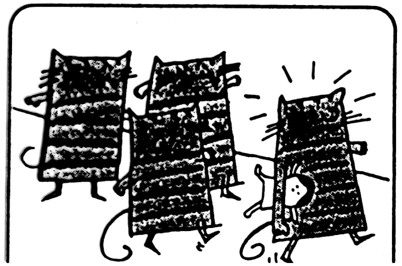
Pri računanju verjetnosti se je lahko zmotiti. Pred nami sta maček in mačka, ki skupaj preživljata čas.



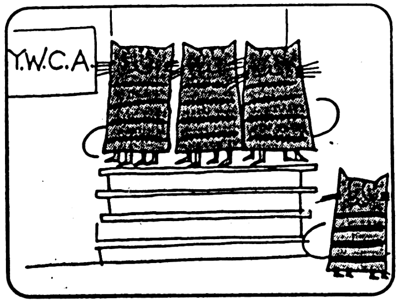
Maček: Ljubica, koliko muck se je skotilo v zadnjem leglu?
Mačka: Kaj ne znaš šteti? Štiri vendar.
Maček: In koliko je muckov?
Mačka: To je težko ugotoviti. Zdaj še ne vem.



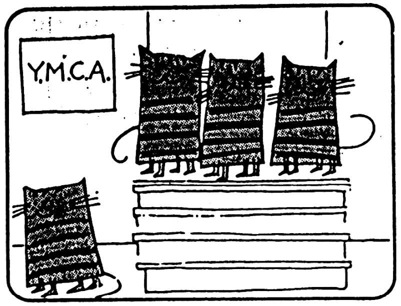
Maček: Ni zelo verjetno, da so vsi štirje mladiči mucki.



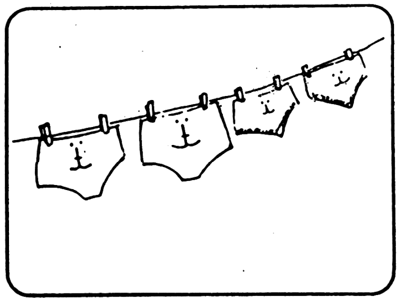
Mačka: In ni verjetno, da so vse mucke.



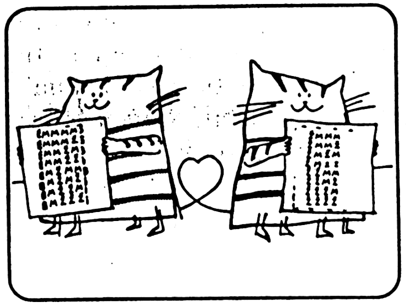
Maček: Mogoče je samo en mucek.



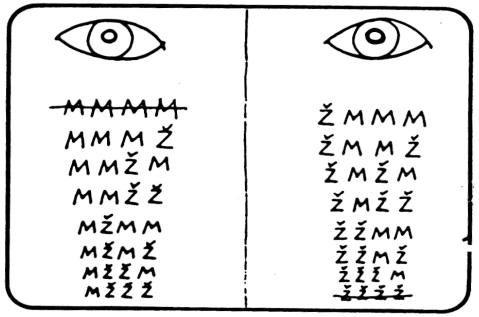
Mačka: Mogoče je pa samo ena mucka.



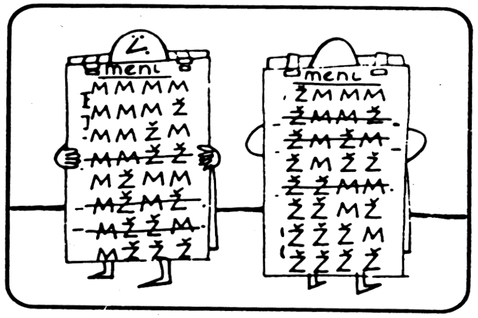
Maček: Saj ni tako težko izračunati. Obstaja 50-odstotna verjetnost, da je vsak mladič mucek ali mucka. Torej je jasno, da imava najverjetneje dva mucka in dve mucki. Ali si jim že dala imena?



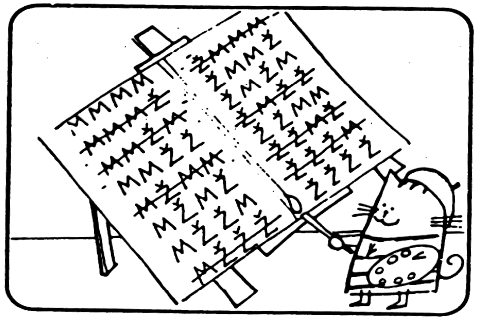
Ali je maček pravilno sklepal? Pa preverimo njegovo teorijo. Označimo mucka z M (moški spol) in mucko z Ž (ženski spol) in napišimo vseh 16 enako verjetnih možnosti.



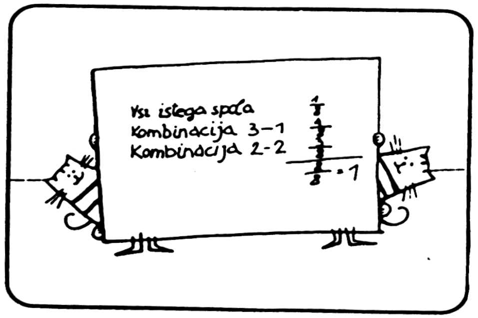
Samo v 2 primerih od 16 so vsi mladiči istega spola. Torej je verjetnost, da se to zgodi, $\frac{2}{16}$ ali $\frac{1}{8}$. Maček je imel prav, ko je dejal, da je tak izid malo verjeten.



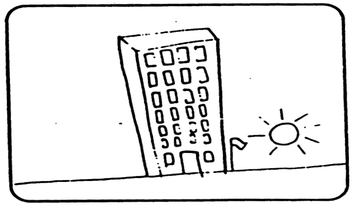
Zdaj pa poglejmo, kolikšna je verjetnost, da se rodita dva mucka in dve mucki, kar se je zdelo mačku najverjetneje. To se zgodi šestkrat, torej je verjetnost $\frac{6}{16}$ ali $\frac{3}{8}$, kar je gotovo več kot $\frac{1}{8}$. Morda ima maček prav.



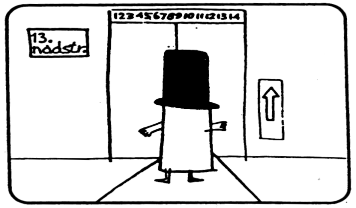
Ogledati pa si moramo še eno možnost: rojstvo 3 muckov in ene mucke ali 3 muck in enega mucka. Ker se to zgodi v 8 primerih, je verjetnost $\frac{8}{16}$ ali $\frac{1}{2}$, kar je celo več, kot pri kombinaciji 2–2, tj. dveh muckov in dveh muck. Ali je možno, da smo se zmotili?



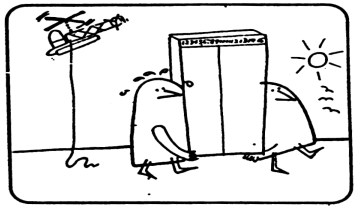
Če smo vse verjetnosti pravilno izračunali, mora biti vsota 1. In res je. Gotovo se bo zgodila ena od treh kombinacij. Maček se je zmotil: najverjetnejša kombinacija ni 2–2, ampak 3–1.



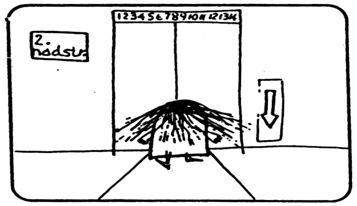
Ljudje, ki se vozijo z dvigalom, se pogosto čudijo še enemu nenavadnemu paradoksu verjetnosti. Recimo, da dvigala v tej zgradbi vozijo neodvisno in da je povprečni čas čakanja v vsakem nadstropju isti.



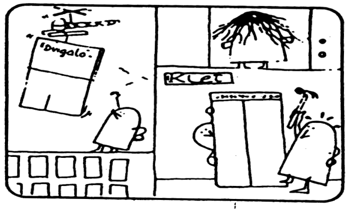
Gorenc ima pisarno v enem od višjih nadstropij. Zelo je nejevoljen.
Gorenc: Presneto! Prvo dvigalo, ki se ustavi tukaj, gre gor. To se dogaja ves čas.



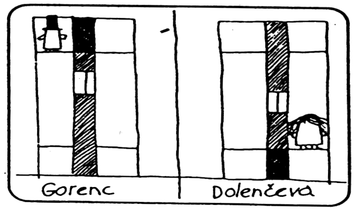
Gorenc: Morda pa dvigala delajo v kleti, s strehe pa jih odnašajo helikopterji.



Dolenčeva dela v enem od nižjih nadstropij. Vsak dan kosi v restavraciji, ki je čisto na vrhu zgradbe. Tudi ona se jezi.
Dolenčeva: Tega sploh ne razumem. Kadarkoli čakam na dvigalo, gre prvo, ki pride, navadno dol.

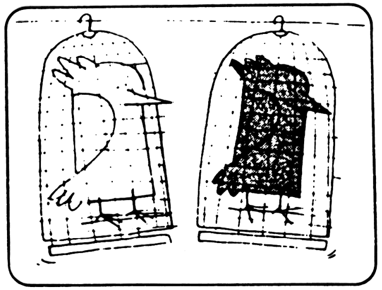


Dolenčeva: Dvigala gotovo nosijo na streho, od tam pa jih spuščajo v skladišče v kleti.

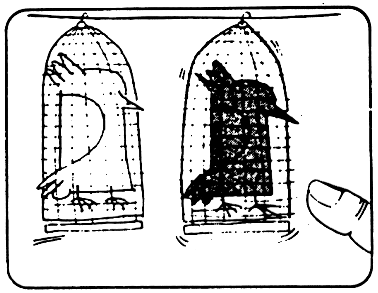


Uganko reši preprost diagram. Za Gorenca gredo dol samo dvigala iz počrnjenega dela jaška. Ta del je majhen, če ga primerjamo s svetlim delom, zato je veliko večja verjetnost, da je dvigalo pod njegovim nadstropjem in da prihaja navzgor. Isto, samo obrnjeno, velja za Dolenčevo.

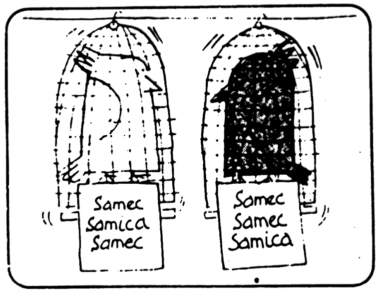
Nenavadna papagaja



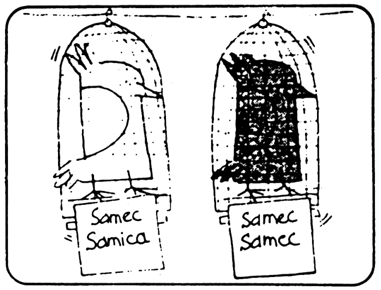
Gospa je imela dva papagaja. Nekega dne jo je vprašal gost:
Gost: Ali je en papagaj samček?
Lastnica: Da, je.
Kolikšna je verjetnost, da sta oba papagaja samčka?
Verjetnost je enaka $\frac{1}{3}$.



Recimo, da gost vpraša:
Gost: Ali je temni papagaj samček?
Lastnica: Da.
Zdaj se verjetnost, da sta oba papagaja samčka, poveča na $\frac{1}{2}$. To je čudno. Zakaj se verjetnost spremeni, če vprašamo za temnega papagaja?



Ta paradoks zlahka razložimo, če napišemo vse možnosti. Če gost ve, da je en ptič samec, so možni samo trije primeri. Samo v enem sta oba papagaja samca, torej je verjetnost, da sta oba papagaja samca, enaka $\frac{1}{3}$. (Predpostavljamo, da je enaka verjetnost, da je vsak papagaj samec ali samica.)



Ko pa gost ve, da je temni ptič samček, sta samo dve možnosti. Samo pri eni sta oba papagaja samca, torej je verjetnost, da sta oba papagaja samca, enaka $\frac{1}{2}$.

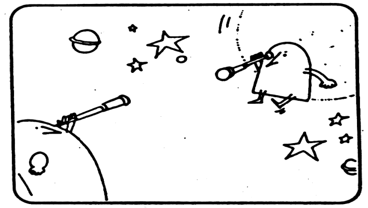
2.3.4.2. Celotna verjetnost

- Celotno verjetnost izračunamo kot povprečne vrednosti pogojnih verjetnosti A, ob upoštevanju verjetnosti B in $\overline{B}$ kot uteži:

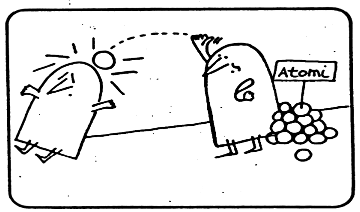
$$Pr(A) = Pr(A|B) \cdot Pr(B) + Pr(A|\overline{B})) \cdot Pr(\overline{B})$$

- Če se A lahko pojavi v katerihkoli medsebojno izključujočih se okoliščinah B_1, B_2, ..., B_k, potem je celotna verjetnost dogodka A:

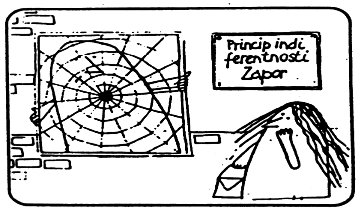
$$Pr(A) = Pr(A|B_1) \cdot Pr(B_1) + Pr(A|B_2) \cdot Pr(B_2) + \ldots + Pr(A|B_k) \cdot Pr(B_k)$$



Ali obstaja življenje na Titanu, največjem Saturnovem satelitu?



Ali bo prišlo do atomske vojne?



Če na taka vprašanja odgovorite, da sta pritrdilni in nikalni odgovor enako verjetna, po neumnem uporabljate »princip indiferentnosti«. Nepremišljena raba tega principa je mnoge matematike, znanstvenike in celo velike filozofe zapeljala v mreže absurdnosti.

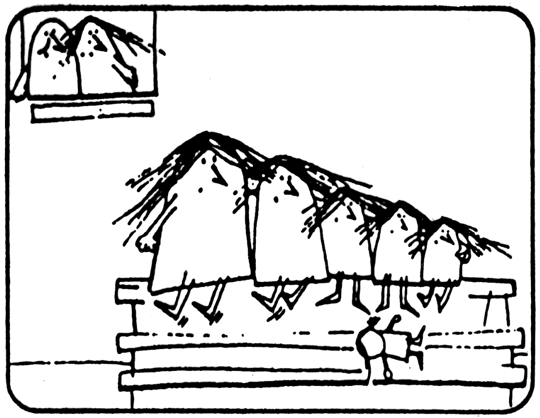
Oglejmo si, kakšna protislovja nastanejo, če ta princip površno uporabimo pri vprašanjih o Titanu in atomski vojni. Kolikšna je verjetnost, da je na Titanu življenje? Uporabimo princip indiferentnosti in dobimo za odgovor $\frac{1}{2}$. Kolikšna je verjetnost, da na Titanu *ni* preprostih rastlin? Odgovor je spet $\frac{1}{2}$. In da ni praživali? Odgovor je spet $\frac{1}{2}$. Kolikšna je verjetnost, da na Titanu ni niti preprostih rastlin niti praživali? Po zakonih verjetnosti moramo pomnožiti $\frac{1}{2}$ z $\frac{1}{2}$ in dobimo $\frac{1}{4}$. To pomeni, da se je verjetnost, da je na Titanu *neke* vrste življenje, povečala na $1 - \frac{1}{4} = \frac{3}{4}$, kar je v nasprotju z našo prejšnjo oceno, ki je bila $\frac{1}{2}$.

Kolikšna je verjetnost atomske vojne pred letom 2000? Po principu indiferentnosti odgovorimo: $\frac{1}{2}$. Kolikšna je verjetnost, da na Združene države Amerike ne bo padla atomska bomba? Odgovor: $\frac{1}{2}$. In da ne bo padla na Sovjetsko zvezo? Odgovor je spet $\frac{1}{2}$. In da ne bo padla na Francijo! Odgovor: $\frac{1}{2}$. Če tako sklepanje uporabimo pri desetih različnih državah, je verjetnost, da atomska bomba ne bo padla na nobeno od njih, $(\frac{1}{2})^{10}$ ali $\frac{1}{1024}$. Če to odštejemo od 1, dobimo verjetnost $\frac{1023}{1024}$, to je verjetnost, da bo atomska bomba padla vsaj na eno od desetih držav.

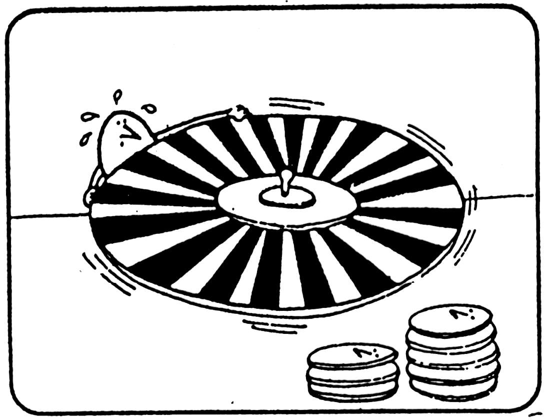
2.3.5. Neodvisnost

- Dogodka A in B sta neodvisna, kadar verjetnost dogodka A ni odvisna od tega, ali se dogodek B zgodi ali ne.
- Uporabimo le Pr(A) namesto Pr(A|B), ker je podatek na desni nebistven.
- Pravilo množenja za izračun verjetnosti več neodvisnih dogodkov je:

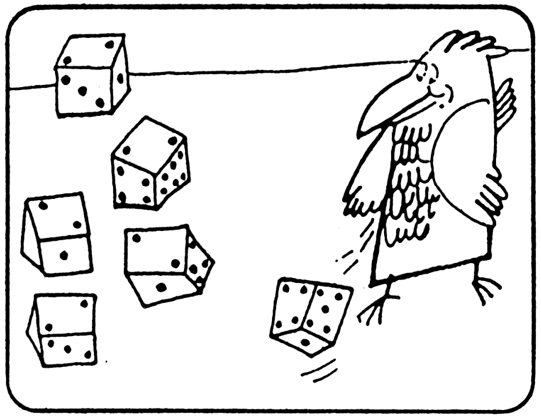
$$Pr(A \text{ in } B \text{ in } C) = Pr(A) \cdot Pr(B) \cdot Pr(C)$$



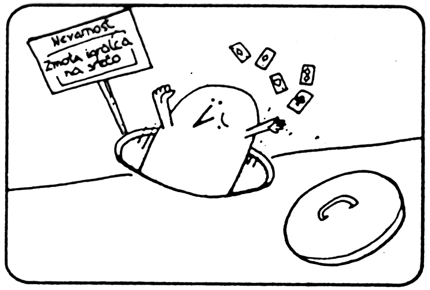
Jerebovi imajo pet otrok – same deklice.
Gospa Jereb: Upam, da naš naslednji otrok ne bo spet deklica.
Gospod Jereb: Draga moja, po petih deklicah bo zdaj gotovo deček.
Ali ima gospod Jereb prav?



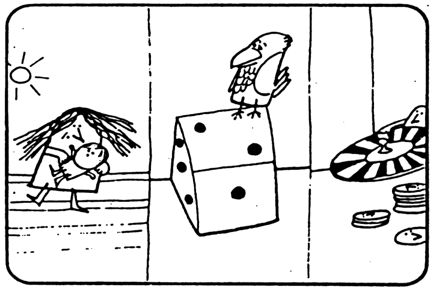
Mnogi igralci mislijo, da bodo zmagali na ruleti, če čakajo, da se večkrat zapovrstjo ustavi na rdečih številkah, potem pa stavijo na črno. Ali bo tak sistem uspešen?



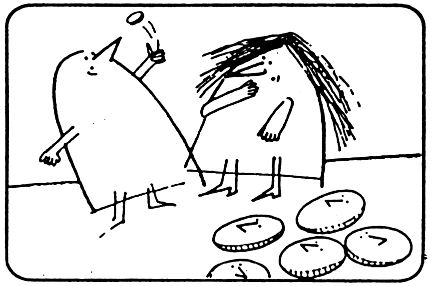
Edgar Allan Poe je dokazoval, da je, če vržeš petkrat zapovrstjo dvojko, verjetnost, da jo boš pri naslednjem metu spet vrgel, manjša od $\frac{1}{6}$. Ali je imel prav?



Če ste na katero od teh vprašanj odgovorili z »da«, ste padli v past, ki jo imenujemo »zmota igralca na srečo«. V vsakem od primerov je naslednji dogodek popolnoma neodvisen od vseh prejšnjih.



Verjetnost, da bo šesti otrok Jerebovih deklica, je enaka kot je bila verjetnost, da bo njihov prvi otrok deklica. Verjetnost, da bo naslednja številka na ruleti rdeča, je enaka kot verjetnost, da je bila prejšnja številka rdeča. Verjetnost, da boste pri naslednjem metu kocke vrgli dvojko, je še vedno $\frac{1}{6}$.



Da vam bo to še bolj jasno, si predstavljajte, da gospod Jereb vrže v zrak 5 kovancev zapovrstjo in vsi padejo tako, da imajo grb na vrhu. Verjetnost, da bo naslednji kovanec vrgel tako, da bo spet imel grb na vrhu, je točno taka kot prej: $\frac{1}{2}$. Kovanec se ne »spomni«, kako je padel pri prejšnih metih.

Če izid dogodka A vpliva na dogodek B, pravimo, da je B »odvisen« od A. Na primer, verjetnost, da boste jutri oblekli dežni plašč, je odvisna od tega, kolikšna je verjetnost, da bo jutri dež ali (točneje), kako vi to verjetnost ocenjujete. Dogodke, za katere v navadnem jeziku pravimo, »da nimajo niti najmanjše zveze drug z drugim«, imenujemo »neodvisne« dogodke. Verjetnost, da boste jutri oblekli dežni plašč, je neodvisna od verjetnosti, da bo imel predsednik države za zajtrk jajca.

2.3.6. Naključne spremenljivke

- Vsak izid poizkusa lahko izrazimo s številčno vrednostjo. Stopnje zaupanja rezulatat ne moremo napovedati, zato pravimo, da je poizkus naključen
- Naključna spremenljivka povezuje numerični rezultat z naključnim poizkusom.

2.3.6. Naključne spremenljivke

POIZKUS	ISKANI IZID	NAKLJUČNA SPREMENLJIVKA
Detekcija radioaktivnosti	Število α delcev	0, 1, 2, ...
Štetje zrn kremena v zbrusku (30 točk)	Število zrn	0, 1, 2, ..., 30
Analiza vsebnosti P_2O_5 v tleh	Odstotek	$0 < P_2O_5 < 100$
Prisotnost fosila v kamnini	Je ali ni	x = O za odsoten x = 1 za prisoten
Uvrsti sediment v pesek, melj ali glino	Vrsta sedimenta	x = 1 za pesek x = 2 za melj x = 3 za glino

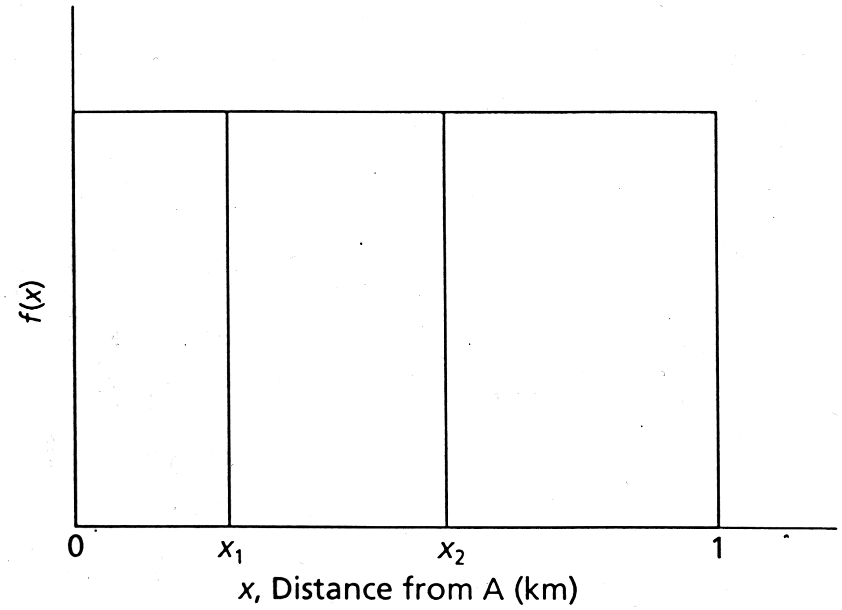
2.3.6. Naključne spremenljivke

- Ne moremo napovedati vrednosti naključne spremenljivke, lahko pa izračunamo verjetnost, da bo opazovanje imelo določeno vrednost oz. vrednost znotraj danega razpona.
- Način, kako je celotna verjetnost 1, porazdeljena med možne vrednosti naključne spremenljivke, imenujemo verjetnostna porazdelitev.

2.3.6. Naključne spremenljivke

- Za zvezno naključno spremenljivko velja

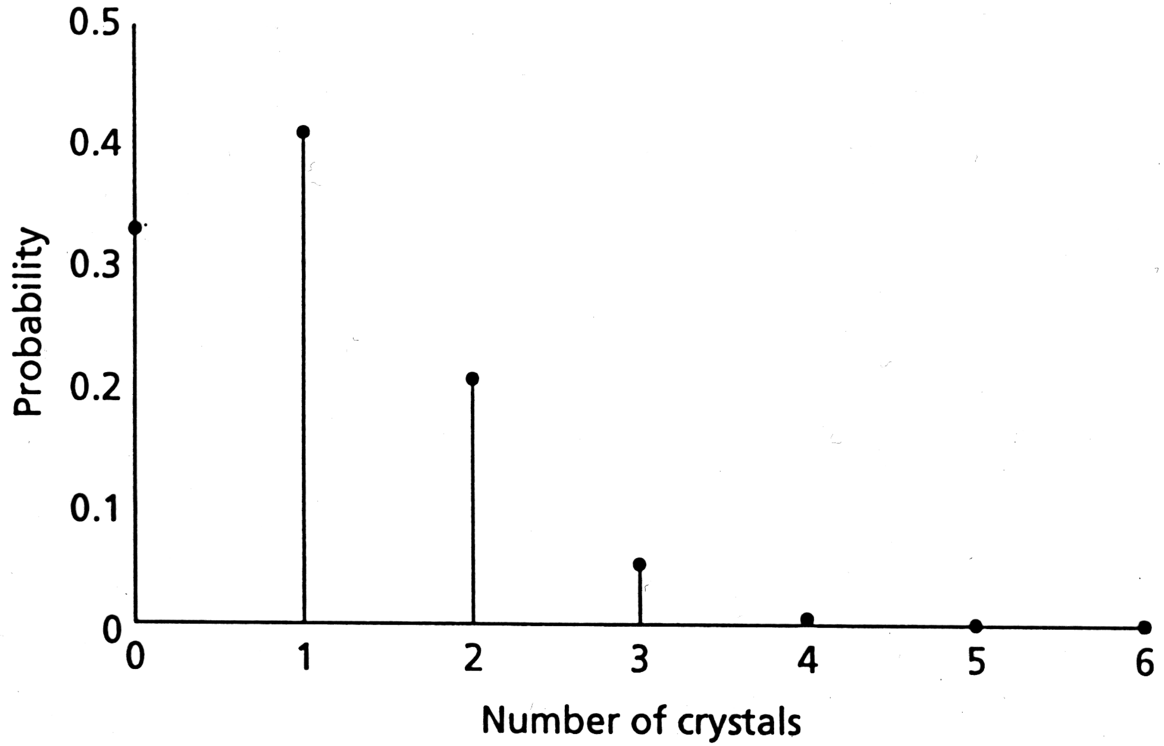
$$\Pr(x_1 < X < x_2) = \frac{(x_2 - x_1)}{1}$$



- Verjetnost podaja površina pod krivuljo med ustreznima mejama.

2.3.6. Naključne spremenljivke

Število opazovanih kristalov	Verjetnost Pr(X=x)
0	0,3277
1	0,4096
2	0,2048
3	0,0512
4	0,0064
5	0,0003
Skupaj	1,0000



2.3.6.1. Pričakovanje naključne spremenljivke

- Zmnožek pozitivnih izidov in njihovih verjetnosti seštejemo v skupni kazalec – pričakovanje E(X) ali povprečje pozitivnega izida v posameznem poizkusu.

$$E(X) = \sum x \cdot \Pr(X = x)$$

- Vsaka opazovana spremenljivka je obtežena s pripadajočo relativno pogostnostjo.

2.3.6.1. Pričakovanje naključne spremenljivke

- Primer:

Verjetnost zaslužka izkoriščanja naftnega polja

Zaslužek (mio EUR)	-10	20	30
VERJETNOST	0.2	0.7	0.1

$$E(X) = -10 \cdot 0{,}2 + 20 \cdot 0{,}7 + 30 \cdot 0{,}1 = 15 \text{ mio EUR}$$

2.3.6.2. Varianca naključne spremenljivke

- Varianca je mera odstopanja naključne spremenljivke od srednje vrednosti. Je srednja vrednost med možnimi vrednostmi naključne spremenljivke in njeno srednjo vrednostjo μ.

$$Var(X) = \sigma_x^2 = E\left\{(X-\mu)^2\right\}$$

- Za nezvezno (diskretno) spremenljivko je:

$$Var(X) = \sigma_x^2 = \sum (X-\mu)^2 \Pr(X = x)$$

- Obtežitev vrednosti, ki se redko pojavljajo, je nizka.

2.3.6.3. Srednja vrednost in varianca preprostih funkcij naključne spremenljivke

- Pogosto želimo izračunati srednje vrednosti in variance baključnih spremenljivk, ki smo jih izpeljali iz druge spremenljivke tako, da:
  - Prištevamo ali množimo naključno spremenljivko s konstanto
  - Med seboj seštevamo ali množimo naključne spremenljivke

2.3.6.3. Srednja vrednost in varianca preprostih funkcij naključne spremenljivke

- Kadar je X naključna spremenljivka, a in c pa konstanti, velja:

$$E(X+a)=E(X)+a$$

$$E(cX)=c\cdot E(x)$$

$$Var(X+a)=Var(X)$$

$$Var(cX)=c^2 Var(X)$$

- Kadar so X_1, ..., X_p naključne spremenljivke, velja

$$E(X_1\pm\ldots\pm X_p)=E(X_1)\pm\ldots\pm E(X_p)$$

- Kadar so spremenljivke nekorelirane, je:

$$Var(X_1\pm\ldots\pm X_p)=Var(X_1)\pm\ldots\pm Var(X_p)$$

- Kadar sta X_1 in X_2 korelirani (povezani):

$$Var(X+Y)=Var(X)+Var(Y)\pm 2\operatorname{cov}(X,Y)$$

2.3.7. Nezvezne verjetnostne porazdelitve

- Verjetnost izida izračunamo s pomočjo verjetnostnih funkcij, ki so različne za različne mehanizme opazovanih procesov.
- Nezvezne
  - Enolična
  - Binomna
  - Poissonova
  - Negativna binomna

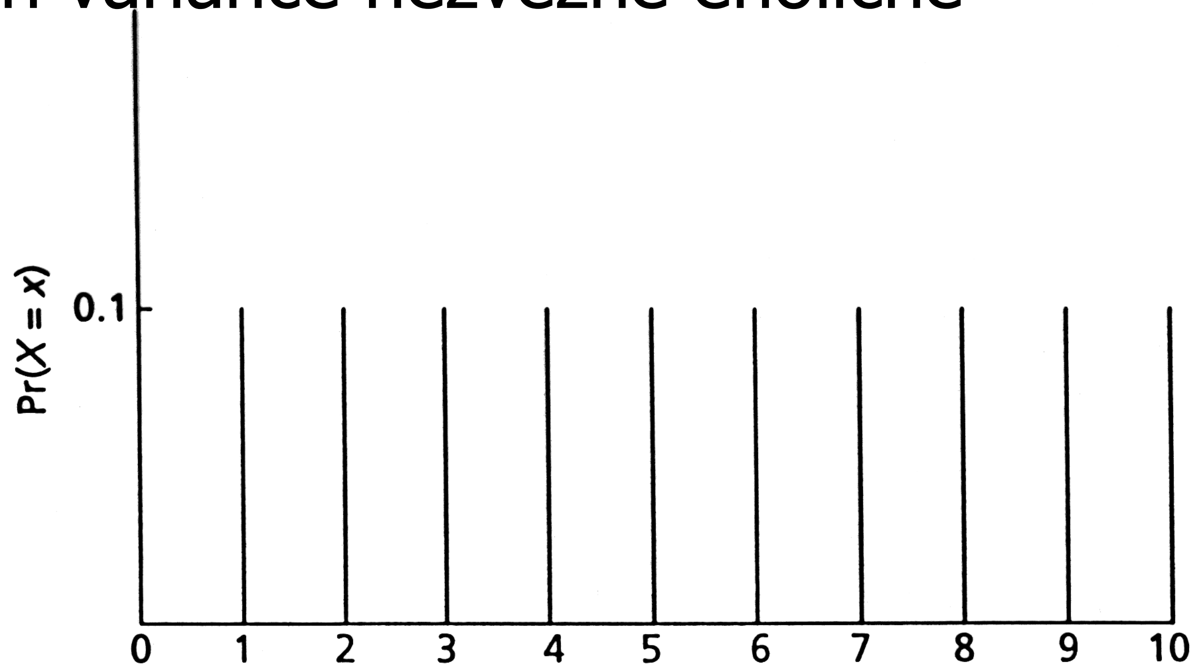
2.3.7.1. Nezvezna enolična porazdelitev

- Naključna spremenljivka, s končnim številom vrednosti, katerih verjetnosti, da se pojavijo so enake, ima nezvezno enoločno (diskretno uniformno) verjetnostno porazdelitev.
- Kadar je n možnih vrednosti X, je verjetnostna funkcija:

$$\Pr(X = x) = \frac{1}{n}$$

2.3.7.1. Nezvezna enolična porazdelitev

- Take porazdelitve v naravi ni. Pomembna je pri naključnem vzorčenju in računalniških modelih, ki skušajo posnemati naravne procese.
- Srednje vrednosti in variance nezvezne enolične verjetnostne porazdelitve so odvisne od možnih vrednosti X.



2.3.7.2. Binomna (dvočlenska) porazdelitev

- Binomno porazdelitev uporabimo, kadar ima poskus le dva možna izida (= Bernoullijev poskus) ter pri neparametrični statistiki.
- Pogoji uporabe binomne porazdelitve:
  - Končno število n neodvisnih Bernoullijevih poskusov
  - Verjetnost obeh izidov je za vsak poskus stalna.
- Iskani rezultat, neodvisna spremenljivka, je število "uspehov" v n poskusih.

2.3.7.2. Binomna (dvočlenska) porazdelitev

- Kadar je $n = 1$, se porazdelitev imenuje Bernoullijeva.
- Če je verjetnost uspeha p, je verjetnost neuspeha $1 - p$.
- Verjetnostno funkcijo zapišemo:

$$\Pr(X=x) = 1 - p \quad \text{za} \quad X = 0$$

$$\Pr(X=x) = p \quad \text{za} \quad X = 1$$

2.3.7.2. Binomna (dvočlenska) porazdelitev

- Srednja vrednost porazdelitve je:

$$E(X) = 0 \cdot (1 - p) + 1p = p$$

- In varianca:

$$Var(X) = (0 - p)^2(1 - p) + (1 - p)^2 p = p(1 - p)$$

- Rezultate lahko uporabimo, kadar ima n visoko vrednost – število poskusov je veliko.

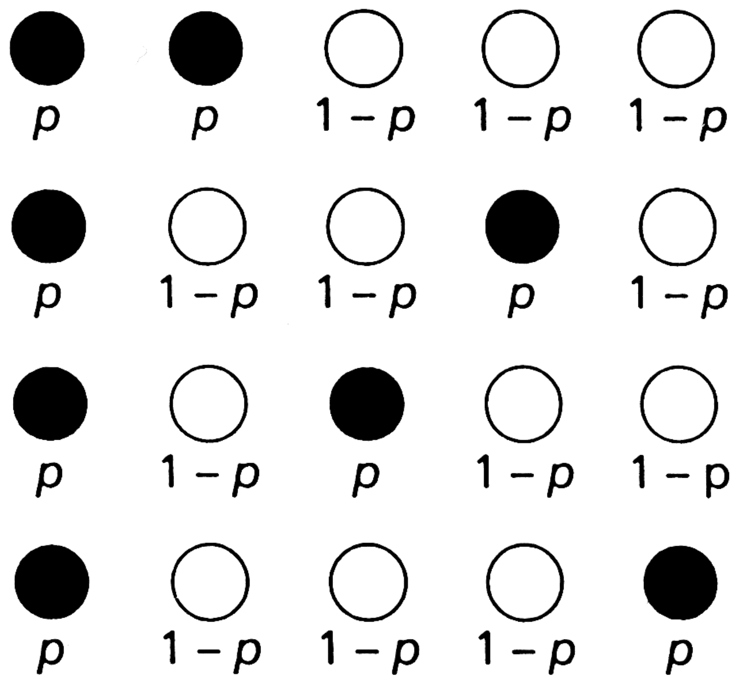
2.3.7.2. Binomna (dvočlenska) porazdelitev

- Število ureditve n objektov, kadar je r enakih ter preostalih (n – r) prav tako enakih, je:

$$\binom{n}{r} = \frac{n(n-1)\cdot\ldots\cdot(n-r+1)}{1\cdot 2\cdot\ldots\cdot r}$$

$$\Pr(X = x) = \binom{n}{x} p^{x}(1-p)^{n-x}$$

$$\mathrm{x} = 0, 1, \ldots, \mathrm{n}$$



2.3.7.2. Binomna (dvočlenska) porazdelitev

- Navadno nas zanima verjetnost števila uspehov znotraj nekega razpona, npr., da bo vrednost višja ali manjša ali enaka neki vrednosti. Seštejemo verjetnosti opazovanja vrednosti v danem razponu:

- $Pr(X>x) = Pr(X=x+1)+Pr(X=x+2)+...+Pr(X=n)$
- $Pr(X\leq x) = Pr(X=x)+Pr(X=x-1)+...+Pr(X=0)$

- Druga od vsot je porazdelitvena funkcija $F(x)$.

		p				
n	*x*	0.1	0.2	0.25	0.3	0.5
2	0	0.810	0.640	0.563	0.444	0.250
	1	0.180	0.320	0.375	0.444	0.500
	2	0.010	0.040	0.063	0.111	0.250
3	0	0.729	0.512	0.422	0.296	0.125
	1	0.243	0.384	0.422	0.444	0.375
	2	0.027	0.096	0.141	0.222	0.375
	3	0.001	0.008	0.016	0.037	0.125
4	0	0.656	0.410	0.316	0.198	0.063
	1	0.292	0.410	0.422	0.395	0.250
	2	0.049	0.154	0.211	0.296	0.375
	3	0.004	0.026	0.047	0.099	0.250
	4	0.000	0.002	0.004	0.012	0.063
5	0	0.590	0.328	0.237	0.132	0.031
	1	0.328	0.410	0.396	0.329	0.156
	2	0.073	0.205	0.264	0.329	0.313
	3	0.008	0.051	0.088	0.165	0.313
	4	0.000	0.006	0.015	0.041	0.156
	5	0.000	0.000	0.001	0.004	0.031
10	0	0.349	0.107	0.056	0.017	0.001
	1	0.387	0.268	0.188	0.087	0.010
	2	0.194	0.302	0.282	0.195	0.044
	3	0.057	0.201	0.250	0.260	0.117
	4	0.011	0.088	0.146	0.228	0.205
	5	0.001	0.026	0.058	0.137	0.246
	6	0.000	0.006	0.016	0.057	0.205
	7	0.000	0.001	0.003	0.016	0.117
	8	0.000	0.000	0.000	0.003	0.044
	9	0.000	0.000	0.000	0.000	0.010
	10	0.000	0.000	0.000	0.000	0.001

2.3.7.2. Binomna (dvočlenska) porazdelitev

- Binomna porazdelitev, kjer je število poskusov n in verjetnost uspeha v vsakem poskusu p, ima srednjo vrednost:

  $\mu = n \cdot p$

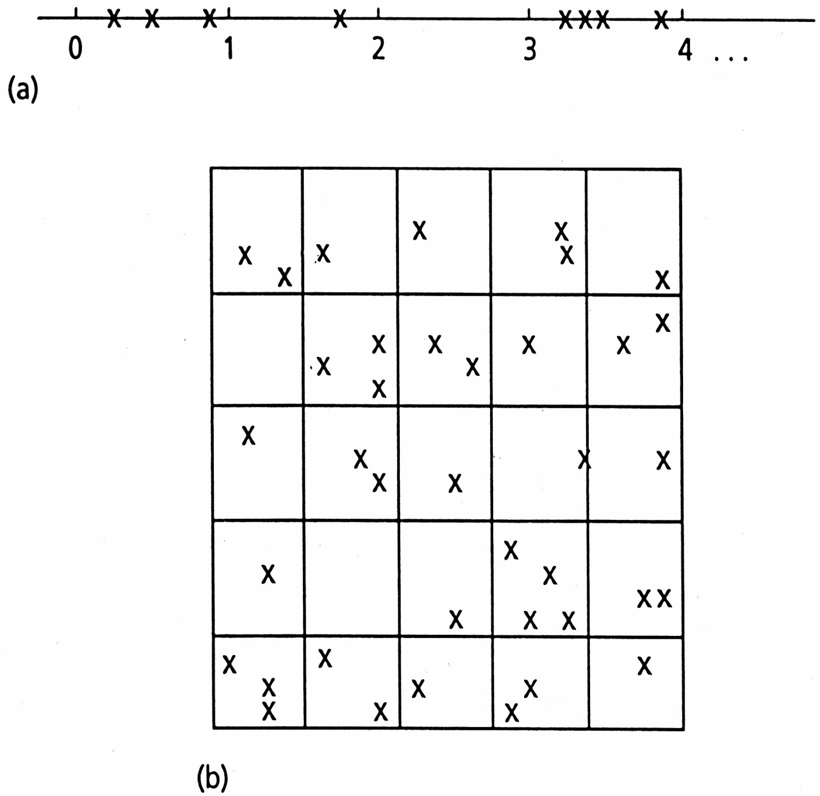
- In varianco:

  $\sigma^2 = n \cdot p \cdot (1 - p)$

- Spremenljivost v številu uspehov je največja, kadar je p = 0,5. To je pomembno kot vodilo pri določanju zahtevanega števila poskusov, potrebnega za oceno deleža primerkov z dano lastnostjo.

2.3.7.3. Poissonova porazdelitev

- Veliko naravnih pojavov obravnavamo kot točke v času ali prostoru.
- Njihovo proučevanje zahteva štetje dogodkov v danem razponu časa ali prostora.



2.3.7.3. Poissonova porazdelitev

- Zanima nas, ali se dogodki pojavljajo naključno in neodvisno.
- Poissonova porazdelitev je model za tako naključnost.
- Dogodki so naključni in neodvisni, kadar velja:
  - Verjetnost posameznega dogodka, ki se pojavi v kratkem časovnem ali prostorskem intervalu, je približno sorazmerna dolžini intervala.
  - Verjetnost, da se v takem intervalu pojavi več kot en dogodek, je dejansko nič.
  - V neprekrivajočih se intervalih je pojavljanje ali nepojavljanje dogodkov neodvisno.

2.3.7.3. Poissonova porazdelitev

- Kadar navedena pravila držijo, ima število dogodkov, ki se pojavijo v končnem intervalu *t* enot časa ali prostora Poissonovo porazdelitev z verjetnostno funkcijo:

$$\Pr(X = x) = \exp(-\mu t) \cdot (\mu t)^x / x! \qquad x = 0, 1, \ldots$$

Kjer je μ povprečno število dogodkov v enoti intervala.

2.3.7.3. Poissonova porazdelitev

	μ			
x	0.5	1.0	1.5	2.0
0	0.607	0.368	0.223	0.135
1	0.303	0.368	0.335	0.271
2	0.076	0.184	0.251	0.271
3	0.013	0.061	0.126	0.180
4	0.002	0.015	0.047	0.090
5	0.000	0.003	0.014	0.036
6	0.000	0.001	0.004	0.012
7	0.000	0.000	0.001	0.003
8	0.000	0.000	0.000	0.001
9	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000

2.3.7.3. Poissonova porazdelitev

- Primer:

  V 50-letnem obdobju se povprečno pojavita 2,2 poplavi. Kakšne so verjetnosti:

  I. Točno dveh poplav v 50 letih?
  II. Točno ene poplave v 25 letih?
  III. Najmanj ene poplave v 50 letih?
  IV. Ne več kot dveh poplav v 25 letih?

2.3.7.3. Poissonova porazdelitev

- Časovna enota, v kateri je bilo ugotovljeno povprečje je 50 let, torej je za 50-letno obdobje t=1, za 25 letno pa polvico, torej t=0,5. Srednja vrednost μ = 2,2.

$$\Pr(X = x) = \exp(-\mu t) \cdot (\mu t)^x / x!$$

I. Verjetnost točno dveh poplav v 50 letih je:

$$\Pr(X = 2) = \exp(-2{,}2 \cdot 1) \cdot \frac{(2{,}2 \cdot 1)^2}{2!} = 0{,}268$$

II. Verjetnost točno ene poplave v 25 letih je:

$$\Pr(X = 1) = \exp(-2{,}2 \cdot 0{,}5) \cdot \frac{(2{,}2 \cdot 0{,}5)^1}{1!} = 0{,}366$$

2.3.7.3. Poissonova porazdelitev

III. Verjetnost najmanj ene poplave v 50 letih je vsota verjetnosti 1, 2, 3,...poplav in jo je lažje izračunati kot komplementaren dogodek 1 – nič poplav:

$$\Pr(X=0)=\exp(-2{,}2\cdot 1)\frac{(2{,}2\cdot 1)^0}{0!}=0{,}1108$$

$$\Pr(X\geq 1)=1-0{,}1108=0{,}8892$$

IV. Verjetnost ne več kot dveh poplav v 25 letih je vsota verjetnosti 0, 1 in 2 poplav:

$$\Pr(X\leq 2)=\Pr(X=0)+\Pr(X=1)+\Pr(X=2)=0{,}9$$

2.3.7.4. Negativna binomna porazdelitev

- Negativna binomna porazdelitev je porazdelitev najmanjšega števila poskusov, potrebnega za določeno število uspehov.
  - Izvajamo neodvisne Bernoullijeve poskuse do prvega uspeha, pri čemer je verjetnost uspeha v kateremkoli poskusu p.
  - Verjetnostna funkcija za število poskusov so vključno prvega uspešnega je:

    $$\Pr(X=x) = (1 - p)^{x-1}p \qquad x = 1, 2, \ldots$$

  - Zaporedne vrednosti napredujejo geometrično, zato se ta, najpreprostejša oblika negativne binomne porazdelitve imenuje tudi geometrična porazdelitev.

2.3.7.4. Negativna binomna porazdelitev

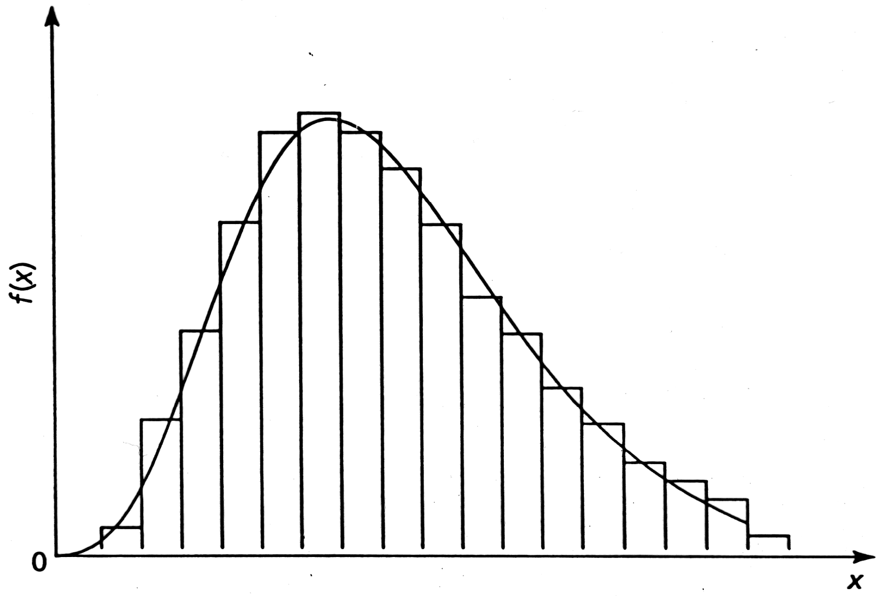
- Negativno binomno porazdelitev uporabimo tudi za porazdelitev števila dogodkov, ki se pojavljajo v intervalu, kadar se srednja vrednost spreminja naključno (= ima eksponencialno ali gama porazdelitev).
- Zanimajo nas verjetnosti števila poskusov, ki bi jih potrebovali, za točno r uspehov.
- Dogodek se bo zgodil na način: $r - 1$ uspehov v $x - 1$ poskusih, pri čemer se r-ti uspeh dogodi v zadnjem, x-tem poskusu.

2.3.7.4. Negativna binomna porazdelitev

- Verjetnost prvega dela dogodka podaja binomna porazdelitev, pri čemer je število poskusov x – 1, nas pa zanima r – 1 uspehov:
$$\binom{r-1}{x-1} = p^{x-1}(1-p)^{(r-1)(x-1)}$$
- Verjetnost drugega dela je preprosto p, in verjetnost obeh dogodkov je zmnožek:
$$\Pr(X=x) = \binom{r-1}{x-1} = p^{x}(1-p)^{r-x}$$
- Srednja vrednost poorazdelitve je: $\frac{r(1-p)}{p}$

2.3.8. Verjetnostne porazdelitve zveznih naključnih spremenljivk

- Če bi bilo število opazovanj izredno veliko ter razredi izredno ozki, bi se zunanja oblika histograma približala zaobljeni krivulji, ki jo imenujemo funkcija gostote verjetnosti – *f(x)*.

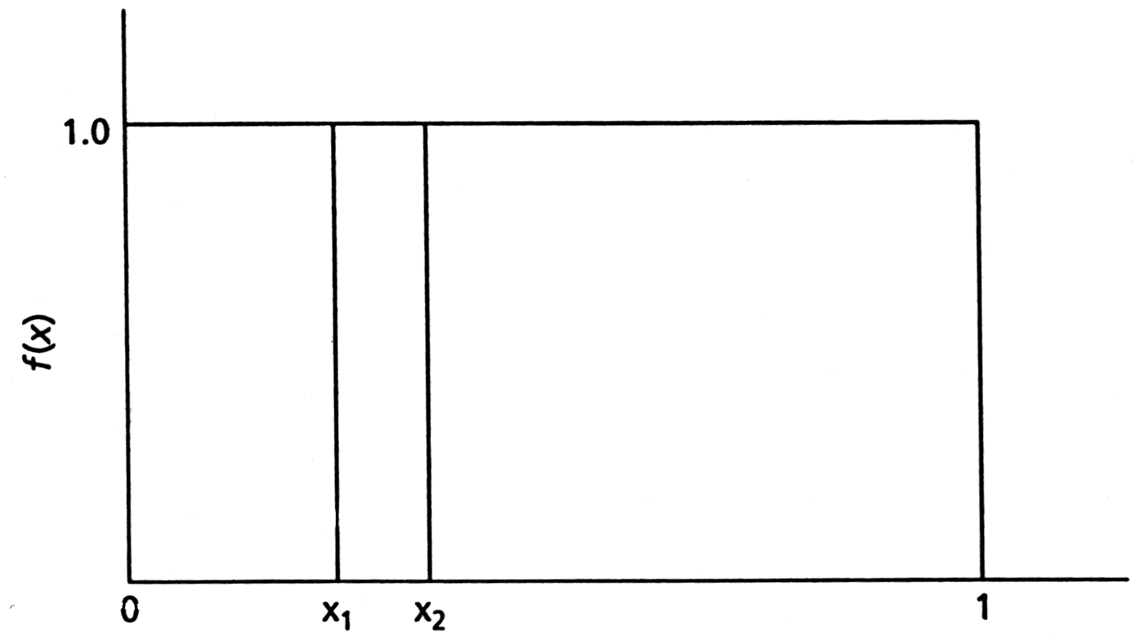


2.3.8. Verjetnostne porazdelitve zveznih naključnih spremenljivk

- Verjetost, da bo opazovanje X ležalo v določenem razponu (x_1, x_2), predstavlja površina pod krivuljo med mejama x_1 in x_2.
- Vrste zvezne porazdelitve:
  - Enolična
  - Eksponenčna
  - Normalna

2.3.8.1. Zvezna enolična porazdelitev

- Naključno spremenljivko s tako porazdelitvijo lahko definiramo v kateremkoli končnem intervalu (0, 1).
- Verjetnost, da je vrednost X v kateremkoli razponu, je sorazmerna njegovi dolžini.



2.3.8.1. Zvezna enolična porazdelitev

- Funkcija gostote verjetnosti (p.d.f.) je:

  $f(x) = 1$ za vse vrednosti X v razponu $0 \leq X \leq 1$

  $f(x) = 0$ za vse ostale možnosti

- Višina p.d.f. je 1 preko celotnega razpona.

  $Pr(x_1 < X \leq x_2) = x_2 - x_1$

  $E(x) = 0{,}5$

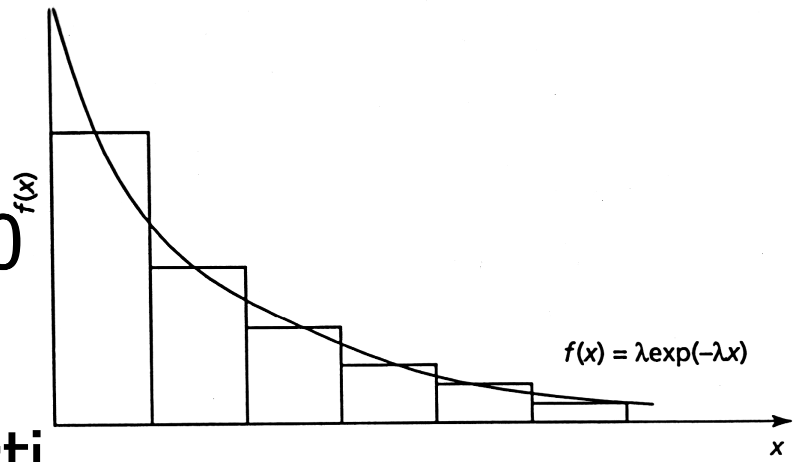
  $Var(X) = 1/12$

2.3.8.2. Eksponenčna porazdelitev

- Kadar se nezvezni dogodki pojavljajo naključno in neodvisno s povprečno stopnjo λ na enoto intervala (število, ki se pojavi v enoti intervala, ima Poissonovo porazdelitev s parametrom λ) je verjetnostna funkcija:

$$f(x) = \lambda \cdot \exp(-\lambda x) \qquad \lambda>0,\ x\geq 0$$

$$0 \qquad \text{za ostale možnosti}$$



2.3.8.2. Eksponenčna porazdelitev

- Povprečna hitrost (stopnja, pogostnost) dogodkov na enoto intervala je λ, zato je povprečen čas med dogodki:

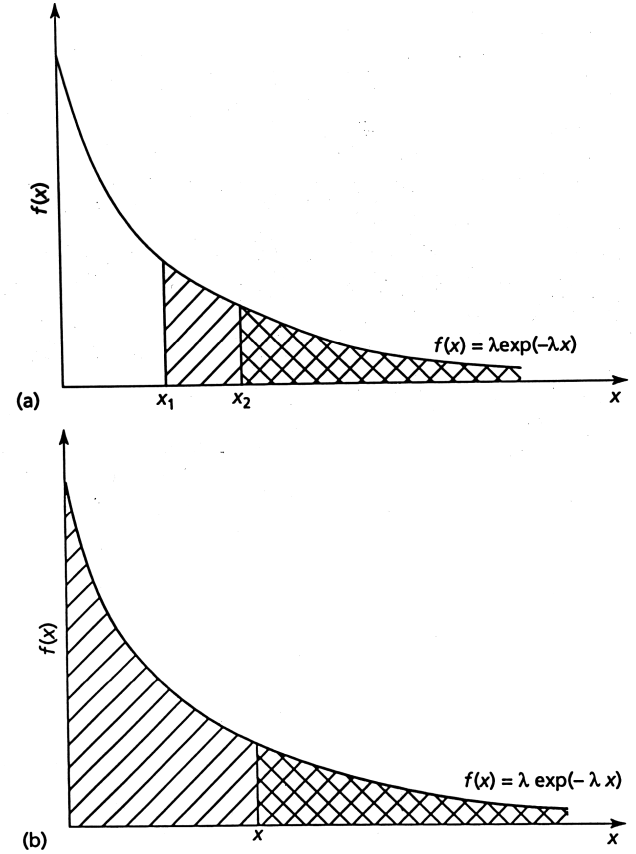
$$E(x) = \frac{1}{\lambda}$$

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

$$\Pr(X \geq x) = \exp(-\lambda x)$$

$$\Pr(x_1 \leq X \leq x_2) = \exp(-\lambda x_1) - \exp(-\lambda x_2)$$

$$\Pr(X \leq x) = 1 - \exp(-\lambda x)$$



2.3.8.2. Eksponenčna porazdelitev

- Primer: Na ljubljanskem območjih sta se v zadnjih nekaj 100-letnih obdobjih pojavila po 2,1 potresa. Kakšna je verjetnost, da je čas med dvema zaporednima potresoma
  I. Preko 25 let
  II. Manj kot 50 let
  III. Med 30 in 40 let

2.3.8.2. Eksponenčna porazdelitev

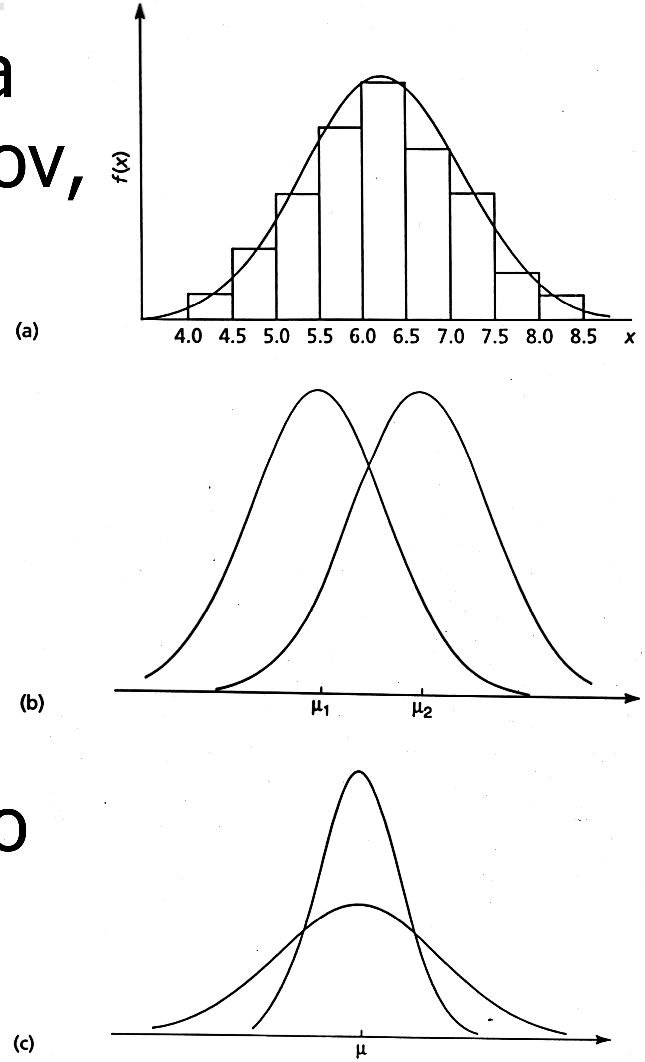
- Ena enota časa T je 100 let, 25 let je torej 0,25 enote, 50 let 0,5...

  $\Pr(T \geq x) = \exp(-\lambda x)$

I. $\Pr(T > 0{,}25) = \exp(-2{,}1 \cdot 0{,}25) = 0{,}59 = 59\%$
II. $\Pr(T < 0{,}5) = 1 - \exp(-2{,}1 \cdot 0{,}5) = 0{,}65 = 65\%$
III. $\Pr(0{,}3 < T < 0{,}4) = \Pr(T > 0{,}3) - \Pr(T > 0{,}4) = \exp(-2{,}1 \cdot 0{,}3) - \exp(-2{,}1 \cdot 0{,}4) = 0{,}1 = 10\%$

2.4. NORMALNA PORAZDELITEV
2.4.1. Pomen normalne porazdelitve

- Kadar so opazovane vrednosti vsota več neodvisnih, naključnih prispevkov, izvirajo iz normalne (Gaussove) porazdelitve, ki je simetrična.
- Prvi parameter funkcije je srednja vrednost, drugi razpon (obseg) porazdelitve na merilni lestvici.
- Vrednosti porazdelitve so lahko pozitivne ali negativne. Matematično popolna normalna porazdelitev se v obe smeri nadaljuje v neskončnost.

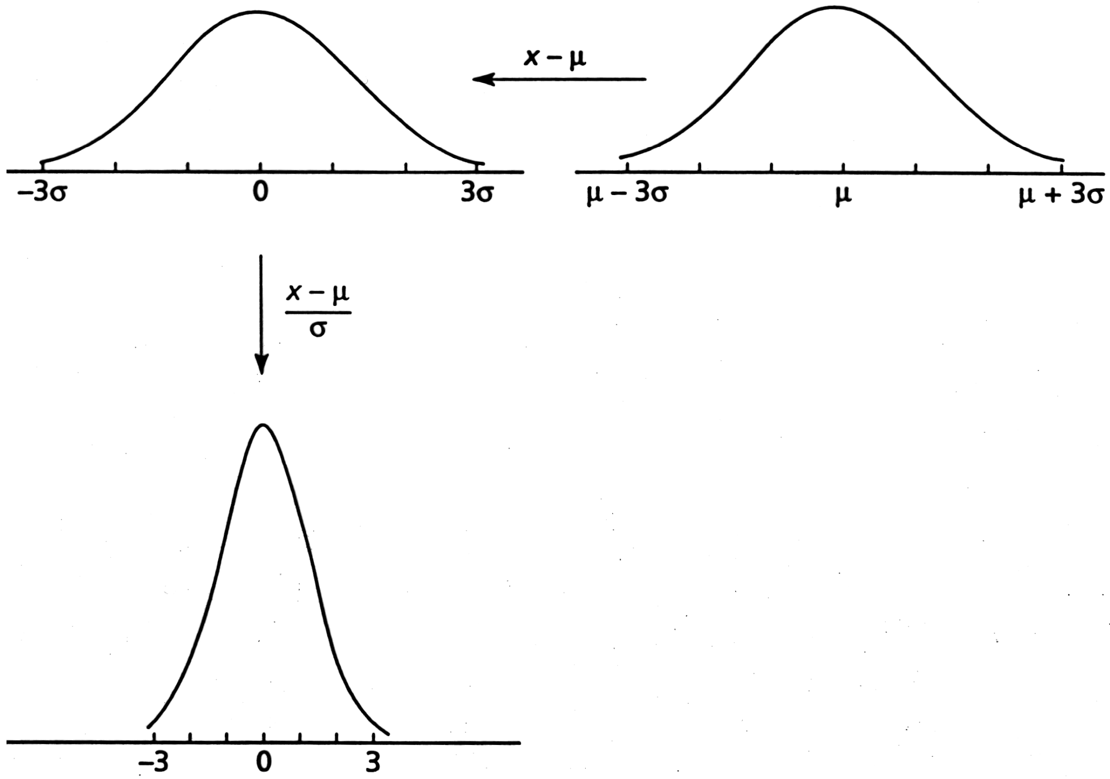


2.4.2. Uporaba tabel normalne porazdelitve

- Izračun ustrezne tabele za vsako možno normalno porazdelitev ni smiselen, zato vrednosti spremenljivke X standariziramo tako, da ustrezajo standardni normalni porazdelitvi s srednjo vrednostjo 0 in standardnim odklonom 1.

2.4.2. Uporaba tabel normalne porazdelitve

- Naključno spremenljivko s srednjo vrednostjo μ in standardnim odklonom σ transformiramo:



$$z = \frac{X - \mu}{\sigma}$$

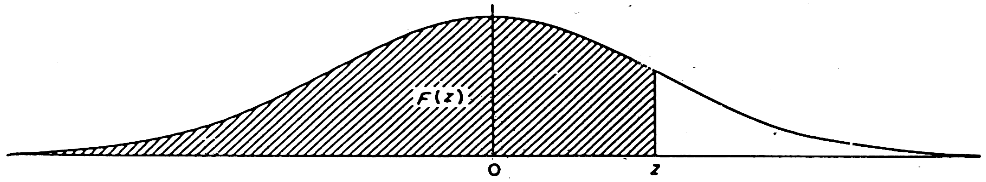


TABLE 1. Cumulative Normal Distribution *

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

* Use explained in Sec. 1.2.1. For more extensive tables, see National Bureau of Standards, *Tables of Normal Probability Functions*, Washington, U. S. Government Printing Office, 1953 (Applied Mathematics Series 23). Note that they show

$$\int_{-z}^{z} f(z)\, dz, \quad \text{not} \quad \int_{-\infty}^{z} f(z)\, dz$$

2.4.2. Uporaba tabel normalne porazdelitve

- Primer: Povprečna vsebnost Pb v Mežici je 12%, standardni odklon je 1,6%. Kakšna je verjetnost, da bo imel naključno izbran kos vsebnsot Pb:

I. 15% ali manj
II. 14% ali več
III. 8% ali manj
IV. Med 8 in 15%?

2.4.2. Uporaba tabel normalne porazdelitve

I. Standarizirana vrednost za $x = 15$ je

$z = (x - \mu)/\sigma = (15-12)/1{,}6 = 1{,}88$

Iz tabele odčitamo, da je ustrezna verjetnost

$Pr(X<15) = Pr(z<1{,}88) = 0{,}9699 = 97\%$

II. Standarizirana vrednost za $x = 14$ je

$z = (x - \mu)/\sigma = (14-12)/1{,}6 = 1{,}25$

in verjetnost zanjo 0,8944. Ker tabela podaja verjetnosti za "manj od". Za "več od" uporabimo komplementarni dogodek $1 - Pr(X<x)$:

$Pr(X\geq 14) = 1 - Pr(X<14) = 1 - 0{,}8944 = 0{,}10565 = 10{,}5\%$

2.4.2. Uporaba tabel normalne porazdelitve

III. Standardizirana vrednost za 8 je

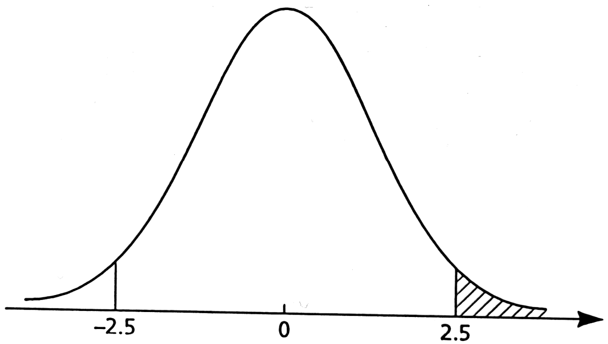
$z = (x - \mu)/\sigma = (8-12)/1,6 = -2,5$

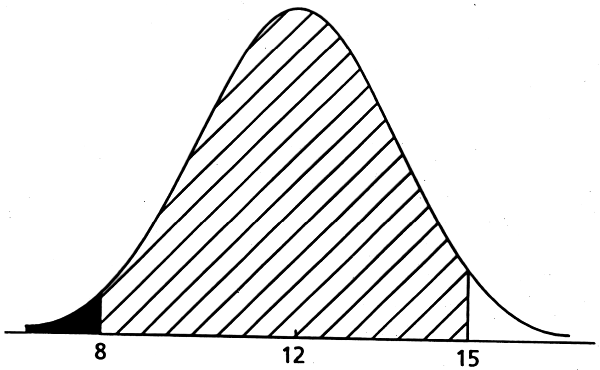
Tabela navaja le pozitivne vrednosti, zato problem rešimo z uporabo simetričnosti normalne porazdelitve $Pr(z \leq -2,5) = Pr(z \geq +2,5)$ in uporabimo enak postopek kot v nalogi II.

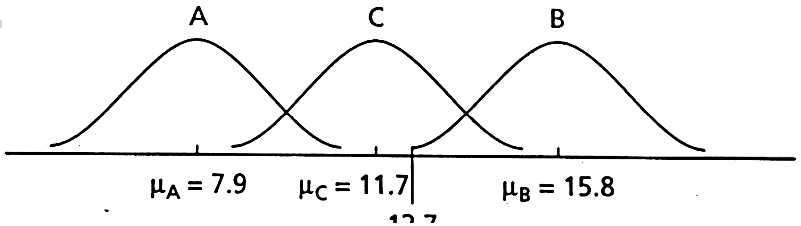
$Pr(x \leq 8) = Pr(z \leq -2,5) = Pr(z \geq 2,5) =$

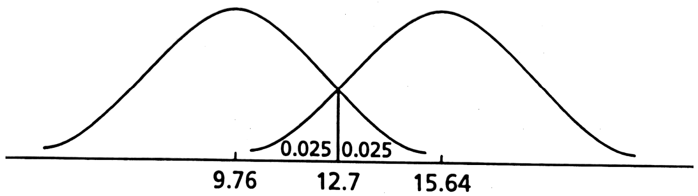
$1 - Pr(z \leq 2,5) = 1 - 0,9938 = 0,00621 = 0,6\%$

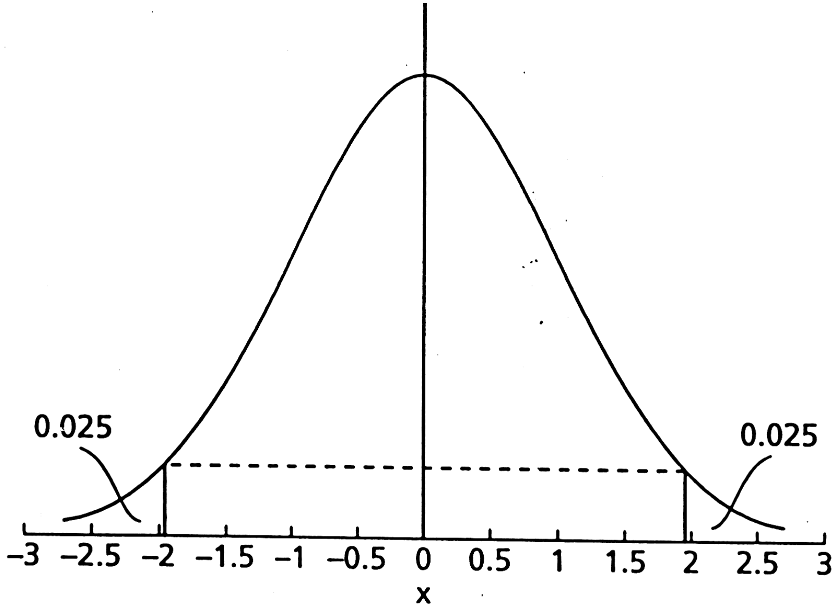
2.4.2. Uporaba tabel normalne porazdelitve











2.4.2. Uporaba tabel normalne porazdelitve

IV. Verjetnost vsebnosti med 8 in 15%, je razlika med verjetnostjo $\leq$15% in $\leq$8%.

$Pr(8<x<15) = Pr(x\leq15) - Pr(x\leq8)=$

$= 0{,}96995 - 0{,}00621 = 0{,}96374 = 96{,}4\%$

2.4.3. Porazdelitev standardizirane normalne spremenljivke

- Kadar ima X normalno porazdelitev, s srednjo vrednostjo μ in standardnim odklonom σ, tedaj ima tudi standardizirana vrednost normalno porazdelitev, s srednjo vrednostjo 0 in standardnim odklonom 1.
- Katerakoli linearna funkcija normalne spremenljivke X, $aX + b$, kjer sta a in b konstanti, ima prav tako normalno porazdelitev; srednja vrednost nove spremenljivke je $a\mu + b$ in standardni odklon $a\sigma$.

2.4.4. Posnemovalne lastnosti normalne porazdelitve

- Pogosto uporabljamo kazalce, ki so linearna kombinacija prvotnih spremenljivk. Npr.:

  $a_1 \cdot$dolžina + $a_2 \cdot$višina + $a_3 \cdot$širina
- Kadar je porazdelitev teh spremenljivk normalna, s srednjimi vrednostmi μ_1, μ_2 in μ_3 ter standardnimi odkloni σ_1, σ_2 in σ_3, tedaj je tudi porazdelitev linearne kombinacije teh spremenljivk normalna.

2.4.4. Posnemovalne lastnosti normalne porazdelitve

- Kadar spremenljivke med seboj niso poovezane (nekorelirane) so kovariance enake 0. Varianca kombinacije je vsota varianc tehtanih spremenljivk.
- Če so X_1, X_2, ..., X_p nekorelirane naključne spremenljivke s srednjimi vrednostmi μ_1, μ_2,..., μ_p ter standardnimi odkloni σ_1, σ_2,..., σ_p, tedaj ima linearna kombinacija

$$y = a_1X_1 + a_2X_2 +...+ a_pX_p$$

normalno porazdelitev s srednjo vrednostjo

$$\mu_y = a_1 \mu_1 + a_2 \mu_2 +...+ a_p \mu_p$$

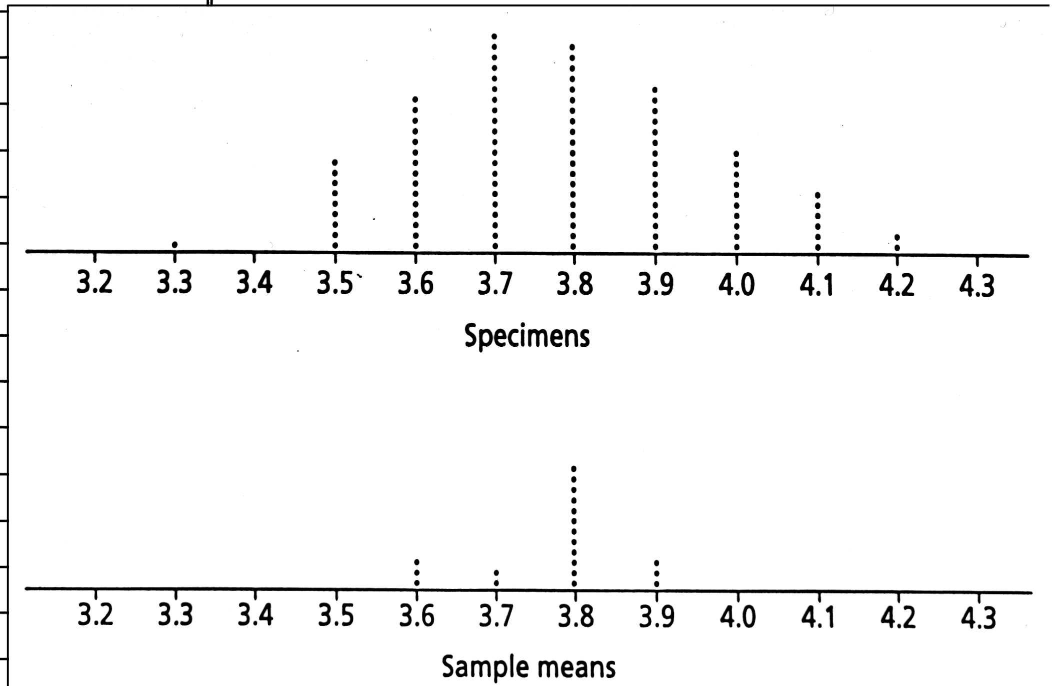
2.4.5. Porazdelitev srednje vrednosti vzorca

- Vsak naključni vzorec bo dal drugačne ocene (statistike) parametrov populacije.
- Statistike izhajajo iz verjetnostne porazdelitve, imenovane vzorčevalna porazdelitev, katere naravo moramo poznati, da lahko sklepamo o parametru, ki nas zanima.

2.4.5. Porazdelitev srednje vrednosti vzorca

Dolžina primerka (cm)

paleontolog	1	2	3	4	5	srednja vrednost vzorca
1	3.6	3.8	3.9	4.1	3.9	3.8
2	3.6	3.9	3.7	3.8	3.9	3.8
3	4.0	4.1	3.8	4.1	3.7	3.9
4	3.7	3.8	3.9	3.9	3.6	3.8
5	3.7	3.7	3.5	4.1	3.7	3.7
6	3.5	3.8	3.6	3.7	3.5	3.6
7	3.9	3.6	4.0	3.6	4.0	3.8
8	3.3	3.6	3.6	3.6	4.0	3.6
9	4.0	3.9	3.5	3.8	3.5	3.7
10	3.8	4.0	3.5	3.9	3.6	3.8
11	3.8	3.7	4.0	3.8	3.7	3.8
12	3.9	3.8	3.7	3.7	3.7	3.8
13	4.0	3.8	3.9	3.8	4.0	3.9
14	3.8	3.9	3.6	3.8	3.7	3.8
15	3.6	3.8	3.7	3.8	4.1	3.8
16	3.5	3.5	3.7	3.7	3.6	3.6
17	3.8	3.5	3.8	3.9	3.8	3.8
18	3.7	3.9	4.2	3.6	3.6	3.8
19	3.9	3.7	3.8	3.7	4.0	3.8
20	4.2	3.7	4.1	3.9	3.7	3.9



2.4.5. Porazdelitev srednje vrednosti vzorca

- Razprševanje srednjih vrednosti več vzorcev okrog skupne srednje vrednosti je manjše, kakor razprševanje vseh opazovanj skupaj.
- Bolj lahko zaupamo srednji vrednosti vzorca več neodvisnih, naključno izbranih primerkov (ali vzorcev) kot posameznemu primerku (ali vzorcu).

2.4.5.1. Srednja vrednost in varianca porazdelitve vzorčenja srednje vrednosti vzorca

- Porazdelitev srednje vrednosti vzorca ima enako srednjo vrednost kot posamezne vrednosti, varainca pa je nižja za faktor, ki je odvisen od velikosti vzorca *n*:

$$\mu_{\bar{x}} = \mu_x$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n}$$

2.4.5.1. Srednja vrednost in varianca porazdelitve vzorčenja srednje vrednosti vzorca

- Kvadratni koren variance vzorčevalne porazdelitve je standardna napaka (*SE* – standard error) statistike:

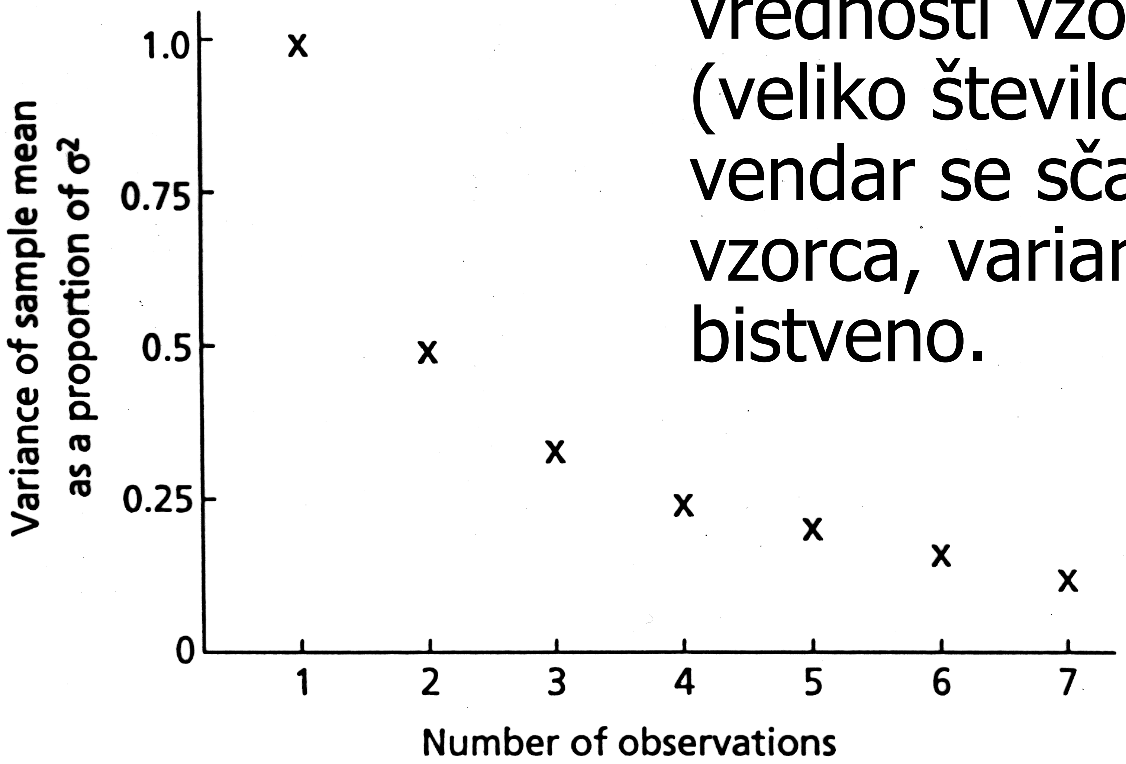
$$SE(\overline{x}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

- Srednje vrednosti vzorca standariziramo podobno kot posamezno vrednost:

$$z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

2.4.5.1. Srednja vrednost in varianca porazdelitve vzorčenja srednje vrednosti vzorca

- Dani populaciji ali izvorni porazdelitvi z varianco σ^2, lahko zmanjšamo varianco srednje vrednosti vzorca z velikim vzorcem (veliko število opazovanj *n*), vendar se sčasoma kljub večanju vzorca, varianca ne zmanjša več bistveno.

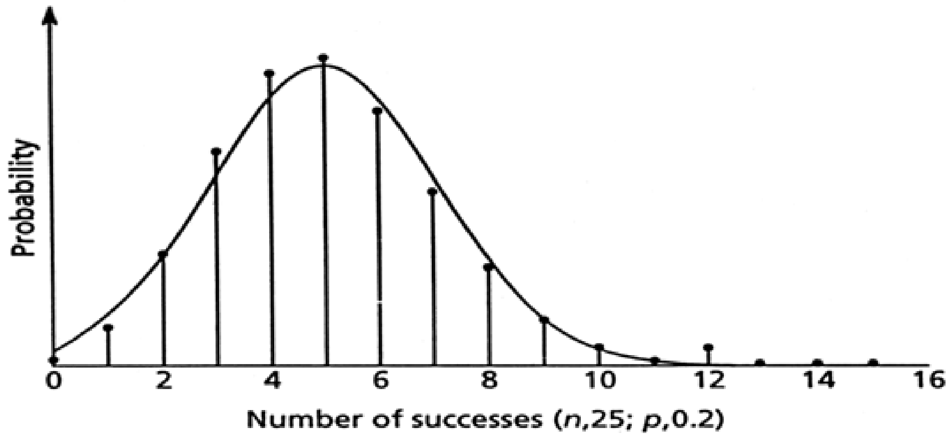


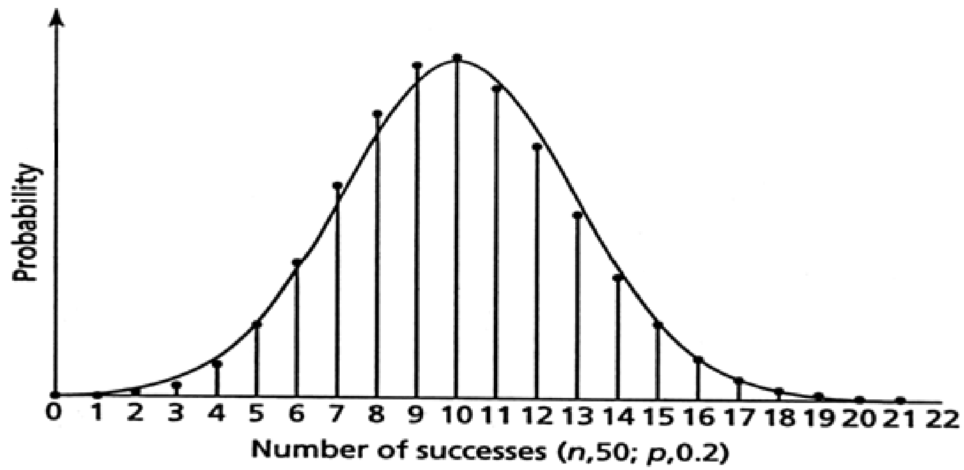
2.4.5.2. Oblike porazdelitve srednjih vrednosti vzorca - teorem približevanja sredini

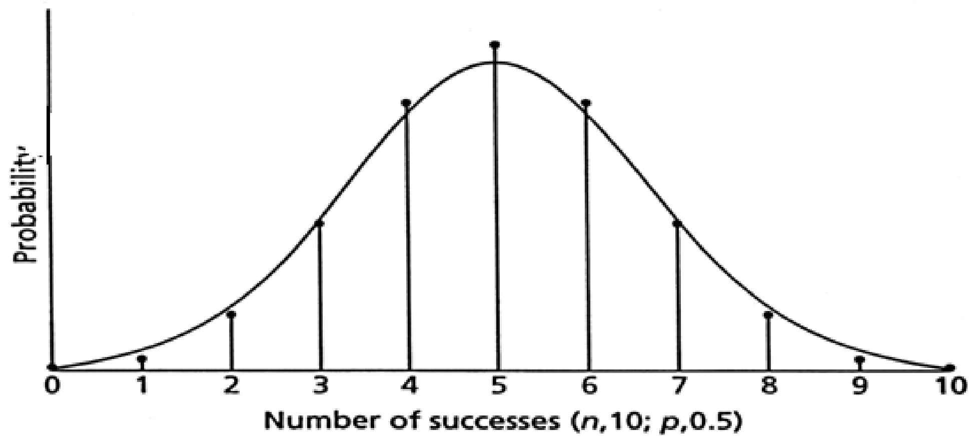
- $X \approx N(\mu, \sigma^2)$ $\qquad \overline{X} \approx N(\mu, \sigma^2/n)$
- Srednje vrednosti velikih vzorcev imajo približno normalno porazdelitev, tudi če dejanska porazdelitev ni taka.
- Približek normalnosti se veča z naraščajočo velikostjo vzorca.
- Če je izvorna porazdelitev simetrična, so celo srednje vrednosti majhnih vzorcev, blizu normalni.

2.4.5.3. Približek normalne porazdelitve binomni

- Kadar verjetnost uspeha v posameznem Bernoullijevem poskusu ni preveč skrajna, ($0,1 \leq p \leq 0,9$), število poskusov pa dovolj veliko, da je $n \cdot p \geq 5$ in $n \cdot (1 - p) \geq 5$, tedaj lahko kot približek binomni porazdelitvi uporabimo normalno, s srednjo vrednostjo $n \cdot p$ in varianco $n \cdot p \cdot (1 - p)$.
- Približek je boljši pri visokih vrednostih n.
- Ujemanje je dobro tudi pri majhnem številu poskusov, kadar je binomna porazdelitev približno simetrična ($p = 0,5$)


Probability
0 2 4 6 8 10 12 14 16
Number of successes (n,25; p,0.2)


Probability
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Number of successes (n,50; p,0.2)


0 1 2 3 4 5 6 7 8 9 10
Number of successes (n,10; p,0.5)

2.4.5.4. Približek normalne porazdelitve Poissonovi

- Kadar ima Poissonova porazdelitev srednjo vrednost $\mu > 30$, se ji lahko približamo z normalno porazdelitvijo s srednjo vrednostjo μ in varianco μ.
- Približek izboljšamo s popravkom zveznosti (±0,5).
- Standarizirana vrednost spremenljivke s Poissonovo porazdelitvijo je tako:

$$z = \frac{x \pm 0{,}5 - \mu}{\sqrt{\mu}}$$