Hermann Þórisson

LÍKINDAREIKNINGUR

YFIRLIT

Kafli	Aðalsíður	hliðarsíður	Myndskeið	klst
LÍKINDI	1-5	1-7	L01-L07	1:12
SLEMBISTÆRÐIR	6-13	14	L08-L14	1:02
VÆNTIGILDI	14-21	13-31	L15-L21	1:20
MIKILVÆGAR STÆRÐIR	22-31		L22-L30	1:14
HÖFUÐ-MARKGILDISREGLAN og POISSONFERLI	32-36	33-36	L31-L35	1:00

Aðalsíður 1-36 eru fyrrihluti þessa heftis.
Á þeim eru dregin saman mikilvægustu atriði efnisins.

hliðarsíður 1-36 eru síðarihluti heftisins. Þær taka við eftir Aðalsíðu 36.
Þar er farið ítarlegar í einstaka atriði.
Á fyrstu hliðarsíðunum er lýst sambandi líkindafræði og tölfræði.

Myndskeiðin L01-L35 byggja á þessu efni.
Ég þakka Júlíu Sif Ólafsdóttur fyrir afgerandi aðstoð við þau.

Prentun: Háskólaprent, 2021
© Hermann Þórisson

LÍKINDI

Yfirlit yfir myndskeið og efni

Skeið	lengd	Aðalsíður	hliðarsíður	efni
L01	9:26		1-2	Inngangur um líkinda- og tölfræði
L02	11:14	A1-A2		Skilgreining og einfaldar reglur
L03	5:42	A3		Jafnar líkur og strjálar
L04	11:03		3-4	Talningarfræði – tengist A3
L05	2:45		5	Dæmi – Jafnar líkur (brids, A3)
L06	14:19	A4	7	Dæmi – Skilyrtar líkur
L07	19:23	A5		Óháðir atburðir

Styttingar í yfirliti eru í samræmi við eftirfarandi dæmi:

L01	fyrsta myndskeið um líkindareikning
9:26	9 mínútur og 26 sekúndur
A1	fyrsta Aðalsíða
1	fyrsta hliðarsíða

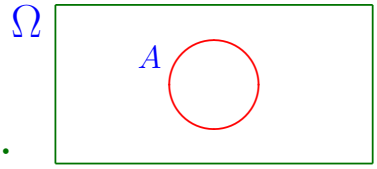
Líkindi – Skilgreining

Hugsum okkur að tilraun sé gerð og setjum

Ω = *útkomumengi* = mengi allra hugsanlegra útkoma

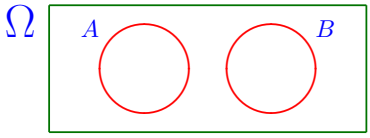
T.D. EF TENINGI ER KASTAÐ: $\Omega = \{1, 2, 3, 4, 5, 6\}$. EÐA LENGD ER MÆLD: $\Omega = [0, \infty)$.

Atburður er hlutmengi í Ω.
Ω kallast *öruggi* atburðurinn.
$\varnothing$ kallast *ómögulegi* atburðurinn.

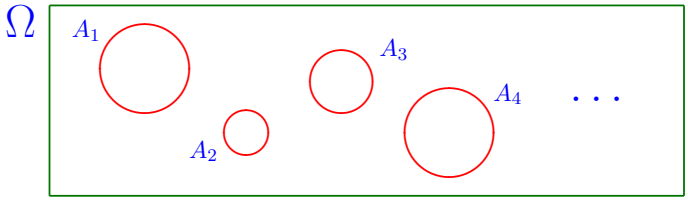


T.D. EF TENINGI ER KASTAÐ ER $\{1, 3, 5\}$ ATBURÐURINN AÐ ODDATALA KOMI UPP.

Atburðir A og B kallast *sundurlægir* ef $A \cap B = \varnothing$.



Atburðir $A_1, A_2, \ldots$ kallast *sundurlægir* ef þeir eru sundurlægir tveir og tveir: $A_i \cap A_j = \varnothing, \ i \neq j$.



Skilgreining: *Líkindi* (eða *líkur*) eru fall $\mathbf{P}$ sem úthlutar sérhverjum atburði A tölu $\mathbf{P}(A)$ þannig að

(1) $\mathbf{P}(\Omega) = 1$.

(2) $\mathbf{P}(A) \geqslant 0$ fyrir alla atburði A.

(3) Ef $A_1, A_2, \ldots$ eru sundurlægir atburðir gildir:

$$\mathbf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbf{P}(A_i).$$

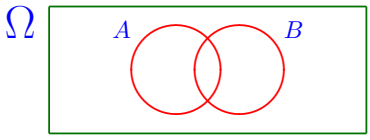
T.D. EF TENINGI ER KASTAÐ KALLAST $\mathbf{P}$ JAFNAR LÍKUR EF $\mathbf{P}(A) = (\text{fjöldi staka í } A)/6$

Nokkrar einfaldar reiknireglur fyrir líkindi

(1) $\mathbf{P}(\varnothing) = 0$ og $0 \leqslant \mathbf{P} \leqslant 1$ og $\mathbf{P}(\Omega) = 1$

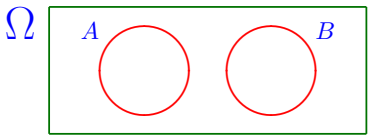
(2) Um alla atburði A og B gildir:

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$$



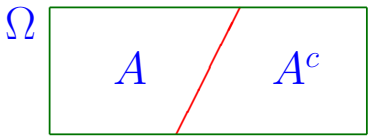
(3) Um sundurlæga atburði A og B gildir:

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$$



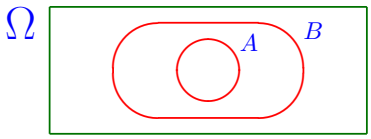
(4) Um alla atburði A gildir:

$$\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$$



(5) Um alla atburði A og B gildir:

$$A \subseteq B \;\Rightarrow\; \mathbf{P}(A) \leqslant \mathbf{P}(B)$$



Jafnar líkur

Skoðum $\Omega = \{\omega_1, \ldots, \omega_n\}$ og setjum $p_i = \mathbf{P}(\{\omega_i\})$.

Skilgreining: Líkindi $\mathbf{P}$ á $\{\omega_1, \ldots, \omega_n\}$ kallast ***jöfn*** ef

$$p_1 = p_2 = \ldots = p_n = \frac{1}{n}.$$

Við segjum þá líka að útkoman sé valin af ***handahófi*** úr menginu $\{\omega_1, \ldots, \omega_n\}$.

Takið eftir að líkindi eru jöfn ef og aðeins ef

$$\mathbf{P}(A) = \frac{\text{fjöldi staka í } A}{n}, \quad A \subseteq \{\omega_1, \ldots, \omega_n\}.$$

Strjál líkindi

Líkindi $\mathbf{P}$ kallast ***strjál*** ef

$$\Omega = \{\omega_1, \ldots, \omega_n\} \quad \text{eða} \quad \Omega = \{\omega_1, \omega_2, \ldots\}.$$

Líkur á atburði $A \subseteq \Omega$ er þá hægt að reikna svona

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\{\omega\}).$$

Takið eftir því að þegar $\Omega = \{\omega_1, \omega_2, \ldots\}$ þá eru jafnar líkur ekki til. Vegna þess að $p_1 = p_2 = p_3 = \ldots$ hefur í för með sér þá mótsögn að $1 \neq 1$,

$$1 = \mathbf{P}(\Omega) = \sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} p_1 = \begin{cases} \infty, & \text{ef } p_1 > 0, \\ 0, & \text{ef } p_1 = 0. \end{cases}$$

Skilyrtar líkur

Skilgreining: Ef A og B eru atburðir og $\mathbf{P}(B) > 0$ þá kallast

$$\mathbf{P}(A\,|\,B) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

skilyrtu líkurnar á A gefið B.

Takið eftir að $\mathbf{P}(A \cap B) = \mathbf{P}(B)\mathbf{P}(A\,|\,B)$.

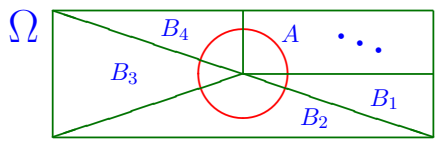
Takið líka eftir að fyrir jafnar líkur gildir (ef $B \neq \varnothing$)

$$\mathbf{P}(A\,|\,B) = \frac{\text{fjöldinn í } A \cap B}{\text{fjöldinn í } B}\,.$$

Sundurlægir atburðir $B_1, B_2, \ldots$ mynda *skiptingu* á Ω ef þeir þekja Ω þ.e. ef $\bigcup_{i=1}^{\infty} B_i = \Omega$.

Skiptingarregla: Ef $B_1, B_2, \ldots$ mynda skiptingu á Ω gildir $\mathbf{P}(A) = \sum_i \mathbf{P}(A \cap B_i)$ og því

$$\mathbf{P}(A) = \sum_i \mathbf{P}(B_i)\mathbf{P}(A\,|\,B_i).$$



Bayes-lögmálið: Ef $\mathbf{P}(A) > 0$ og $\mathbf{P}(B) > 0$ þá gildir

$$\mathbf{P}(B\,|\,A) = \frac{\mathbf{P}(B)\mathbf{P}(A\,|\,B)}{\mathbf{P}(A)}$$

Óháðir atburðir

Það væri eðlilegt að segja að A sé óháður B ef

$$\mathbf{P}(A\,|\,B) = \mathbf{P}(A) \quad \text{þ.e.} \quad \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \mathbf{P}(A)$$

og að B sé óháður A ef

$$\mathbf{P}(B\,|\,A) = \mathbf{P}(B) \quad \text{þ.e.} \quad \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} = \mathbf{P}(B).$$

Þar sem þetta eru sömu eiginleikarnir er óþarfi að taka fram hvor atburðurinn sé óháður hinum.

Skilgreining: Atburðir A og B kallast *óháðir* ef

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B).$$

Skilgreining: Atburðir A, B og C kallast *óháðir* ef

$$\begin{aligned} \mathbf{P}(A \cap B) &= \mathbf{P}(A)\mathbf{P}(B) \\ \mathbf{P}(A \cap C) &= \mathbf{P}(A)\mathbf{P}(C) \\ \mathbf{P}(B \cap C) &= \mathbf{P}(B)\mathbf{P}(C) \\ \text{og} \quad \mathbf{P}(A \cap B \cap C) &= \mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C). \end{aligned}$$

Almenn skilgreining: Atburðir $A_1, \ldots, A_n$ kallast *óháðir* ef fyrir öll k og $n_1 < \ldots < n_k$ gildir:

$$\mathbf{P}(A_{n_1} \cap \ldots \cap A_{n_k}) = \mathbf{P}(A_{n_1}) \cdots \mathbf{P}(A_{n_k}).$$

Atburðir $A_1, A_2 \ldots$ kallast líka *óháðir* ef þetta gildir fyrir öll n.

SLEMBISTÆRÐIR

Yfirlit yfir myndskeið og efni

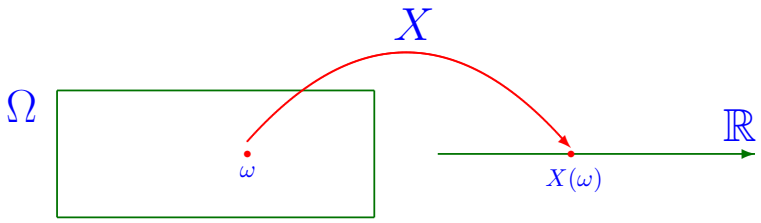
Skeið	lengd	Aðalsíður	hliðarsíður	efni
L08	5:17	A6		Slembistærðir
L09	8:24	A7		Dreififall
L10	8:38	A8		Strjálar stærðir - líkindafall
L11	5:00	A9		Samfelldar stærðir - þéttleiki
L12	10:27	A10-11		Sam-dreififall Sam-líkindafall Sam-þéttleiki
L13	7:01	A12-13		Skilyrðing og óháðar stærðir
L14	17:17		17-19	Dæmi – tengist A11-13

Styttingar í yfirliti eru í samræmi við eftirfarandi dæmi:

L01	fyrsta myndskeið um líkindareikning
9:26	9 mínútur og 26 sekúndur
A1	fyrsta Aðalsíða
1	fyrsta hliðarsíða

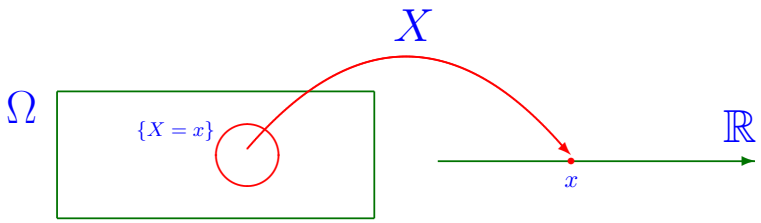
Slembistærðir

Slembistærð **er fall** $X : \Omega \to \mathbb{R}$.



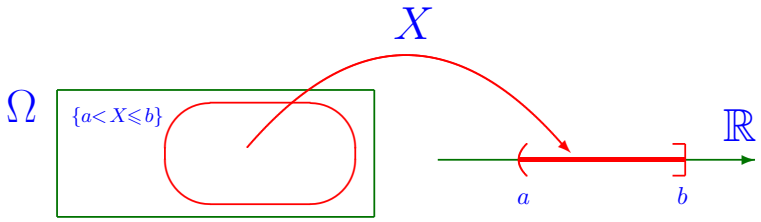
Fyrir $x \in \mathbb{R}$ **skilgreinum við**

$$\{X = x\} := \{\omega \in \Omega : X(\omega) = x\}.$$



Fyrir $a < b$ **skilgreinum við á sama hátt t.d.**

$$\{a < X \leqslant b\} := \{\omega \in \Omega : a < X(\omega) \leqslant b\}.$$



Og almennt fyrir $A \subseteq \mathbb{R}$ **skilgreinum við**

$$\{X \in A\} := \{\omega \in \Omega : X(\omega) \in A\}.$$

Dreififall slembistærðar

Skilgreining: *Dreififall* slembistærðar X er

$$F(x) := \mathbf{P}(X \leqslant x), \quad x \in \mathbb{R}.$$

Skrifum oft F_X í stað F. Þetta er gert til að greina F frá dreififöllum annarra slembistærða.

Regla: Dreififall F uppfyllir:

(1) $0 \leqslant F(x) \leqslant F(y) \leqslant 1, \quad x < y.$

(2) $F(x) \to 0$ þegar $x \to -\infty$.

(3) $F(x) \to 1$ þegar $x \to +\infty$.

(4) F er samfellt frá hægri: $F(y) \to F(x)$ þegar $y \searrow x$.

Aths: Ef fall F hefur þessa 4 eiginleika, þá er til slembistærð X með F sem dreififall.

Regla: $\quad \mathbf{P}(a < X \leqslant b) = F(b) - F(a), \quad a < b.$

Setjum $\quad F(b-) := \lim_{a \nearrow b} F(a), \quad b \in \mathbb{R}.$

Takið eftir að fyrir öll $b \in \mathbb{R}$ gildir

$$\mathbf{P}(X = b) = F(b) - F(b-),$$

og sér í lagi

$$F \text{ samfellt í } b \;\Rightarrow\; \mathbf{P}(X = b) = 0.$$

Strjálar slembistærðir

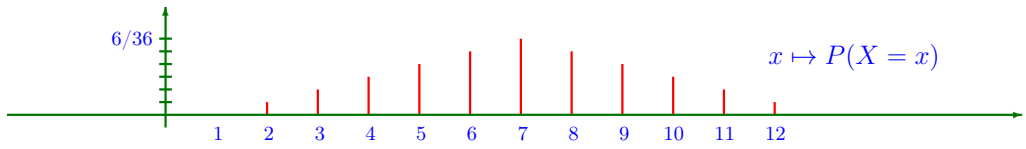
Skilgr: X kallast *strjál* ef hún tekur endanlega mörg gildi $\{a_1, \ldots, a_k\}$ eða teljanlega mörg gildi $\{a_1, a_2, \ldots\}$.

Fallið $x \mapsto \mathbf{P}(X = x)$ kallast þá *líkindafall* X.

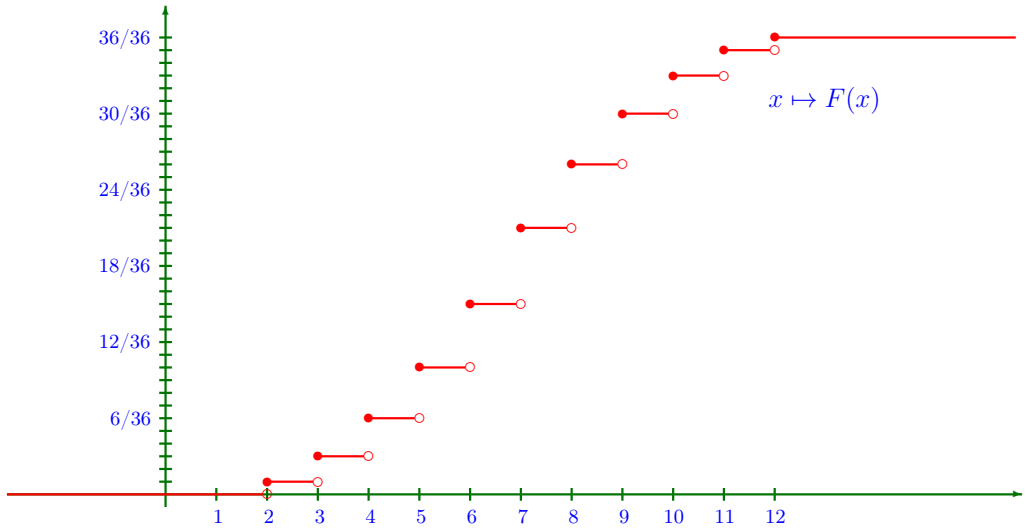
Takið eftir að ef X er strjál þá er dreififallið summa:

$$F(x) := \mathbf{P}(X \leqslant x) = \sum_{y \leqslant x} \mathbf{P}(X = y).$$

Dæmi (tveir teningar): Látum $\mathbf{P}$ vera jafnar líkur á útkomumenginu $\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$. Látum X vera punktasummuna, þ.e. $X(i, j) = i + j$. Líkindafall X er þá svona:



Og dreififall X er svona:

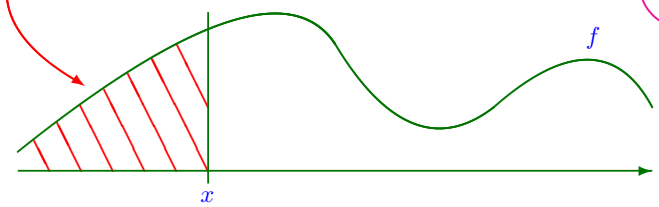


Samfelldar slembistærðir

Skilgreining: Fall f sem uppfyllir $f(x) \geqslant 0$, $x \in \mathbb{R}$, og $\int_{-\infty}^{\infty} f(x)dx = 1$ kallast ***þéttleiki***. Slembistærð X kallast ***samfelld*** ef til er þéttleiki f þ.a.

$$F(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y, \quad x \in \mathbb{R}.$$

$$\boxed{f = F'}$$



f kallast þá þéttleiki X. Skrifum oft f_X í stað f.

Aths: Ef f er þéttleiki, þá er til X með þéttleika f.

Regla: Ef X er samfelld slembistærð, þá er dreififallið F samfellt fall. Ennfremur gildir að

$$\mathbf{P}(a < X \leqslant b) = \int_a^b f(x)\,\mathrm{d}x, \quad a < b,$$

og

$$\mathbf{P}(X = b) = 0, \quad b \in \mathbb{R}.$$

Dæmi: Samfelld slembistærð X kallast ***jöfn*** á bilinu $[a, b]$ (þar sem $a < b$) ef hún hefur þéttleikann

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leqslant x \leqslant b, \\ 0, & \textbf{annars}. \end{cases}$$

Við segjum þá líka að X hafi ***jafna dreifingu*** á $[a, b]$ og að X hafi verið valið ***af handahófi*** á milli a og b.

Samdreiififall

Skoðum slembistærðir $X_1, \ldots, X_n$ sem allar koma úr sömu tilraun, þ.e. sem allar eru skilgreindar á sama útkomumengi, $X_i : \Omega \to \mathbb{R}$, $i = 1, \ldots, n$, með sama $\mathbf{P}$.

Skilgreining: Fallið F sem er skilgreint svona

$$F(x_1, \ldots, x_n) := \mathbf{P}(X_1 \leqslant x_1, \ldots, X_n \leqslant x_n)$$

kallast ***samdreiififall*** slembistærðana $X_1, \ldots, X_n$.

Takið eftir að dreiififall t.d. X_1 fæst svona úr F,

$$F_{X_1}(x) = F(x, \infty, \ldots, \infty)$$

Á sama hátt fást hin dreiififöllin $F_{X_2}, \ldots, F_{X_n}$ úr F. Þessi föll eru oft nefnd ***jaðardreiififöll*** $X_1, \ldots, X_n$.

Samlíkindafall

Skilgreining: Ef $X_1, \ldots, X_n$ eru allar strjálar kallast

$$(x_1, \ldots, x_n) \mapsto \mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

samlíkindafall þeirra.

Takið eftir að líkindafall t.d. X_1 fæst svona:

$$x \mapsto \mathbf{P}(X_1 = x) = \sum_{x_2, \ldots, x_n} \mathbf{P}(X_1 = x, X_2 = x_2, \ldots, X_n = x_n).$$

Á sama hátt fást líkindaföllin fyrir $X_2, \ldots, X_n$. Þessi föll eru oft nefnd ***jaðarlíkindaföll*** $X_1, \ldots, X_n$.

Samþéttleiki

Skilgreining: Látum F vera samdreififall $X_1, \ldots, X_n$. Ef til er fall $f : \mathbb{R}^n \to [0, \infty)$ þannig að

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f(y_1, \ldots, y_n)\, \mathrm{d}y_1 \ldots \mathrm{d}y_n$$

þá kallast f *samþéttleiki* stærðanna $X_1, \ldots, X_n$.

Fallið f er oft fundið með því að diffra

$$F(x_1, \ldots, x_n) = \mathbf{P}(X_1 \leqslant x_1, \ldots, X_n \leqslant x_n)$$

í öllum breytistærðum $x_1, \ldots, x_n$.

Regla: Ef $C \subseteq \mathbb{R}^n$ er t.d. rétthyrningur, hringur, kassi eða kúla, þá má nota f til að reikna líkurnar á því að vigurinn sem stærðirnar mynda lendi í C:

$$\mathbf{P}\big((X_1, \ldots, X_n) \in C\big) = \int \cdots \int_{(x_1, \ldots, x_n) \in C} f(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Regla: Ef $X_1, \ldots, X_n$ hafa samþéttleika, þá eru þær samfelldar hver fyrir sig og þéttleiki t.d. X_1 er

$$f_{X_1}(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x, x_2 \ldots, x_n)\, \mathrm{d}x_2 \ldots \mathrm{d}x_n.$$

Á sama hátt fást hinir þéttleikarnir $f_{X_2}, \ldots, f_{X_n}$ úr f. Þessi föll eru oft nefnd *jaðarþéttleikar* $X_1, \ldots, X_n$.

Aths: Þótt stærðirnar $X_1, \ldots, X_n$ séu samfelldar hver fyrir sig, er ekki víst að þær hafi samþéttleika.

Skilyrt líkindafall

Skilgreining: Ef $X_1, \ldots, X_n$ eru strjálar slembistærðir og $\mathbf{P}(X_2 = x_2, \ldots, X_n = x_n) > 0$, þá kallast fallið

$$x \mapsto \mathbf{P}(X_1 = x \,|\, X_2 = x_2, \ldots, X_n = x_n)$$

skilyrt líkindafall X_1 *gefið* $X_2 = x_2, \ldots, X_n = x_n$.

Munið að $\mathbf{P}(A \,|\, B) := \mathbf{P}(A \cap B)/\mathbf{P}(B)$ og því er

$$\mathbf{P}(X_1 = x \,|\, X_2 = x_2, \ldots, X_n = x_n) := \frac{\mathbf{P}(X_1 = x, X_2 = x_2, \ldots, X_n = x_n)}{\mathbf{P}(X_2 = x_2, \ldots, X_n = x_n)}$$

Skilyrtur þéttleiki

Skilgreining: Látum $X_1, \ldots, X_n$ hafa samþéttleika f. Látum $f_{X_2,\ldots,X_n}$ vera samþéttleika $X_2, \ldots, X_n$. Ef $f_{X_2,\ldots,X_n}(x_2, \ldots, x_n) > 0$, þá kallast fallið

$$x \mapsto f_{X_1|X_2,\ldots,X_n}(x \,|\, x_2, \ldots, x_n) := \frac{f(x, x_2, \ldots, x_n)}{f_{X_2,\ldots,X_n}(x_2, \ldots, x_n)}$$

skilyrtur þéttleiki X_1 *gefið* $X_2 = x_2, \ldots, X_n = x_n$.

Skilgreining: *Skilyrtu líkurnar* á að X_1 taki t.d. gildi á bili A *gefið* $X_2 = x_2, \ldots, X_n = x_n$ eru þá

$$\mathbf{P}(X_1 \in A \,|\, X_2 = x_2, \ldots, X_n = x_n) := \int_A f_{X_1|X_2,\ldots,X_n}(x \,|\, x_2, \ldots, x_n)\, \mathrm{d}x\,.$$

Óháðar slembistærðir

Skilgreining: $X_1, \ldots, X_n$ kallast ***óháðar*** ef

$$\mathbf{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \mathbf{P}(X_1 \in A_1) \ldots \mathbf{P}(X_n \in A_n),$$

fyrir öll bil $A_1, \ldots, A_n$.

Regla: Þetta gildir ef og aðeins ef

$$F(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n), \quad x_1, \ldots, x_n \in \mathbb{R},$$

þar sem F er samdreififallið og F_{X_i} er dreififall X_i:

$$F_{X_i}(x_i) = \mathbf{P}(X_i \leqslant x_i), \quad i = 1, \ldots, n\,.$$

Regla: Strjálar slembistærðir $X_1, \ldots, X_n$ eru óháðar ef og aðeins ef

$$\mathbf{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathbf{P}(X_1 = x_1) \ldots \mathbf{P}(X_n = x_n)$$

fyrir öll $x_1, \ldots, x_n \in \mathbb{R}$.

Regla: Samfelldar slembistærðir $X_1, \ldots, X_n$ eru óháðar ef og aðeins ef

$$(x_1, \ldots, x_n) \mapsto f_{X_1}(x_1) \ldots f_{X_n}(x_n)$$

er samþéttleiki fyrir stærðirnar $X_1, \ldots, X_n$.

VÆNTIGILDI

Yfirlit yfir myndskeið og efni

Skeið	lengd	Aðalsíður	hliðarsíður	efni
L15	9:14	A14	21-23	Væntigildi og miðgildi
L16	11:21	A14-15		Dæmi Lögmál ómeðvitaða tölfræðingsins
L17	13:28	A16	25-30	Reglur um væntigildi
L18	5:37			Dæmið á síðum A14-15, framhald út frá reglunni neðst á A16
L19	13:13	A17	32	Dreifni
L20	15:01	A18	32	Samdreifni og fylgni
L21	12:07	A19-21		Ójöfnur Markovs og Chebyshevs Lögmál mikils fjölda

Styttingar í yfirliti eru í samræmi við eftirfarandi dæmi:

L01	fyrsta myndskeið um líkindareikning
9:26	9 mínútur og 26 sekúndur
A1	fyrsta Aðalsíða
1	fyrsta hliðarsíða

Væntigildi

Skilgreining: ***Væntigildi*** **(*meðalgildi*)** X **er**

$$\mu = \mu_X = \mathbf{E}[X] := \begin{cases} \displaystyle\sum_x x\mathbf{P}(X = x), & \text{ef } X \text{ er strjál,} \\ \displaystyle\int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x, & \text{ef } X \text{ er samfelld.} \end{cases}$$

Dæmi: **Látum** X **hafa þéttleikann**

$$f_X(x) = 1,\ 0 \leqslant x \leqslant 1, \quad (\text{og } f_X(x) = 0 \text{ annars}).$$

Væntigildið fæst svona: $\mathbf{E}[X] = \int_0^1 x\,\mathrm{d}x = \left[\frac{1}{2}x^2\right]_0^1 = \frac{1}{2}$.

Skoðum nú slembistærðina X^3**. Til að finna** $\mathbf{E}[X^3]$ **útfrá skilgreiningunni** $\mathbf{E}[X^3] := \int_{-\infty}^{\infty} y f_{X^3}(y)\,\mathrm{d}y$ **þurfum við að finna** $f_{X^3} = F'_{X^3}$**. Dreififallið** F_{X^3} **fæst svona:**

$$F_{X^3}(y) := \mathbf{P}(X^3 \leqslant y) = \mathbf{P}(X \leqslant y^{\frac{1}{3}}) = \begin{cases} 0, & y < 0, \\ y^{\frac{1}{3}}, & 0 \leqslant y \leqslant 1, \\ 1, & y > 1. \end{cases}$$

Þéttleikinn er því

$$f_{X^3}(y) = F'_{X^3}(y) = \begin{cases} \frac{1}{3}y^{-\frac{2}{3}}, & 0 \leqslant y \leqslant 1, \\ 0, & \text{annars.} \end{cases}$$

Væntigildið fæst nú svona:

$$\mathbf{E}[X^3] := \int_{-\infty}^{\infty} y f_{X^3}(y)\,\mathrm{d}y = \int_0^1 y\,\frac{1}{3}\,y^{-\frac{2}{3}}\,\mathrm{d}y = \frac{1}{3}\left[\frac{y^{\frac{4}{3}}}{\frac{4}{3}}\right]_0^1 = \frac{1}{4}.$$

Takið eftir að $\frac{1}{4} \neq \left(\frac{1}{2}\right)^3$ **þannig að** $\mathbf{E}[X^3] \neq (\mathbf{E}[X])^3$.

Þægileg væntigildisregla

Munið skilgreininguna á væntigildi:

$$\mathbf{E}[X] := \begin{cases} \displaystyle\sum_x x\mathbf{P}(X = x), & \text{ef } X \text{ er strjál,} \\ \displaystyle\int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x, & \text{ef } X \text{ er samfelld.} \end{cases}$$

Lögmál ómeðvitaða tölfræðingsins (ein stærð):
Látum $g : \mathbb{R} \to \mathbb{R}$ vera fall (t.d. $g(x) = x^3$).
Þá gildir að

$$\mathbf{E}[g(X)] = \begin{cases} \displaystyle\sum_x g(x)\mathbf{P}(X = x), & \text{ef } X \text{ er strjál,} \\ \displaystyle\int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x, & \text{ef } X \text{ er samfelld.} \end{cases}$$

Dæmi (framhald): Látum aftur X hafa þéttleikann

$$f_X(x) = 1,\ 0 \leqslant x \leqslant 1, \quad (\text{og } f_X(x) = 0 \text{ annars}).$$

Við fengum áðan eftir vafstur við að finna $f_{X^3}(y)$

$$\mathbf{E}[X^3] := \int_{-\infty}^{\infty} y f_{X^3}(y)\,\mathrm{d}y = \int_0^1 y\,\frac{1}{3}\,y^{-\frac{2}{3}}\,\mathrm{d}y = \frac{1}{3}\left[\frac{y^{\frac{4}{3}}}{\frac{4}{3}}\right]_0^1 = \frac{1}{4}.$$

Við getum nú farið einfaldari leið:

$$\mathbf{E}[X^3] = \int_{-\infty}^{\infty} x^3 f_X(x)\,\mathrm{d}x = \int_0^1 x^3\,\mathrm{d}x = \left[\frac{x^4}{4}\right]_0^1 = \frac{1}{4}.$$

Reglur um væntigildi

Munið skilgreininguna á væntigildi:

$$\mathbf{E}[X] := \begin{cases} \sum_x x\mathbf{P}(X = x), & \text{ef } X \text{ er strjál,} \\ \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x, & \text{ef } X \text{ er samfelld.} \end{cases}$$

Lögmál ómeðvitað tölfræðingsins (margar stærðir):
Látum $g : \mathbb{R}^n \to \mathbb{R}$ vera fall og setjum

$$Y = g(X_1, \ldots, X_n).$$

Þá gildir að

$$\mathbf{E}[Y] = \begin{cases} \sum_{x_n} \cdots \sum_{x_1} g(x_1, \ldots, x_n)\mathbf{P}(X_1 = x_1, \ldots, X_n = x_n) \\ \qquad\qquad \text{ef } X_1, \ldots, X_n \text{ eru strjálar,} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f(x_1, \ldots, x_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \\ \qquad\qquad \text{ef } X_1, \ldots, X_n \text{ hafa samþéttleika } f. \end{cases}$$

Reglur: Fyrir $a, a_1, \ldots, a_n \in \mathbb{R}$ gildir

(1) $\mathbf{E}[a] = a$.

(2) $\mathbf{E}[aX] = a\mathbf{E}[X]$.

(3) $\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y]$.

(4) $\mathbf{E}[a_1X_1 + \ldots + a_nX_n] = a_1\mathbf{E}[X_1] + \ldots + a_n\mathbf{E}[X_n]$

Regla: Ef X og Y eru óháð gildir $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$.

Dreifni

Skilgreining: ***Dreifni*** **slembistærðar** X **er**

$$\mathrm{Var}[X] := \mathbf{E}[(X-\mu)^2] \qquad \big(\textbf{hér er } \mu = \mathbf{E}[X]\big)$$

og ***staðalfrávik*** **hennar er**

$$\sigma := \sqrt{\mathrm{Var}[X]}.$$

Takið eftir að $\mathrm{Var}[X] \geqslant 0$.

Reglur:

(1) $$\mathrm{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2.$$

(2) $$\mathrm{Var}[a+bX] = b^2\mathrm{Var}[X], \quad a, b \in \mathbb{R}.$$

Regla: **Ef** $X_1, \ldots, X_n$ **eru óháðar gildir að**

$$\mathrm{Var}[X_1 + \ldots + X_n] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n].$$

Samdreifni og fylgni

Skilgreining: ***Samdreifni*** **slembistærða** X **og** Y **er**

$$\mathrm{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

og ***fylgnistuðull*** **þeirra er**

$$\rho = \rho_{X,Y} := \frac{\mathrm{Cov}[X,Y]}{\sigma_X \sigma_Y}\,.$$

Takið eftir að $\mathrm{Cov}[X,X] = \mathrm{Var}[X]$.

Regla: $-1 \leqslant \rho \leqslant 1$

Regla: $\mathrm{Cov}[X,Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$

Reglur: Fyrir $a \in \mathbb{R}$ **og** $n, m \in \mathbb{N}$ **gildir að**

(1) $\mathrm{Cov}[aX,Y] = a\,\mathrm{Cov}[X,Y]$

(2) $\mathrm{Cov}[X+Y,Z] = \mathrm{Cov}[X,Z] + \mathrm{Cov}[Y,Z]$

(3) $$\mathrm{Cov}\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}[X_i, Y_j]$$

(4) $$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + 2\sum_{i<j} \mathrm{Cov}[X_i, X_j]$$

Regla: Ef X **og** Y **eru óháðar gildir að** $\mathrm{Cov}[X,Y] = 0$.

Aths: Hins vegar þurfa X **og** Y **ekki að vera óháðar þótt** $\mathrm{Cov}[X,Y] = 0$. **Fylgnileysi útilokar ekki hæði.**

Ójöfnur Markovs og Chebyshevs

Ójafna Markovs: Ef $\mathbf{P}(X \geqslant 0) = 1$ þá gildir að

$$\mathbf{P}(X > a) \leqslant \frac{\mathbf{E}[X]}{a}, \quad a > 0.$$

Sönnun: Setjum

$$Y := \begin{cases} 0, & \text{ef } X \leqslant a, \\ a, & \text{ef } X > a. \end{cases}$$

Þá er $\mathbf{P}(Y \leqslant X) = 1$ sem gefur $\mathbf{E}[Y] \leqslant \mathbf{E}[X]$.

Takið nú eftir að $\mathbf{E}[Y] = a\,\mathbf{P}(X > a)$.

Þetta tvennt gefur að $a\,\mathbf{P}(X > a) \leqslant \mathbf{E}[X]$.

Ójafna Chebyshevs: Ef $|\mathbf{E}[Y]| < \infty$ þá gildir að

$$\mathbf{P}(|Y - \mathbf{E}[Y]| > \varepsilon) \leqslant \frac{\mathrm{Var}[Y]}{\varepsilon^2}, \quad \varepsilon > 0.$$

Sönnun: Setjum $X := (Y - \mathbf{E}[Y])^2$. Þá er

$$\begin{aligned} \mathbf{P}(|Y - \mathbf{E}[Y]| > \varepsilon) &= \mathbf{P}(X > \varepsilon^2) \\ &\leqslant \frac{\mathbf{E}[X]}{\varepsilon^2} \quad \text{(ójafna Markovs)} \\ &= \frac{\mathbf{E}[(Y - \mathbf{E}[Y])^2]}{\varepsilon^2} \\ &= \frac{\mathrm{Var}[Y]}{\varepsilon^2}. \end{aligned}$$

Lögmál mikils fjölda

Munið ójöfnu Chebyshevs: Ef $|\mathbf{E}[Y]| < \infty$ þá gildir að

$$\mathbf{P}(|Y - \mathbf{E}[Y]| > \varepsilon) \leqslant \frac{\mathrm{Var}[Y]}{\varepsilon^2}, \quad \varepsilon > 0.$$

Við notum nú þessa ójöfnu til að sanna merka reglu.

Lögmál mikils fjölda: Ef $X_1, X_2, \ldots$ eru óháðar og hafa allar sama endanlega væntigildi $\mu = \mathbf{E}[X_i]$ og allar sömu endanlegu dreifni $\sigma^2 = \mathrm{Var}[X_i]$, þá gildir fyrir öll $\varepsilon > 0$ að

$$\mathbf{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \varepsilon\right) \to 0, \quad n \to \infty.$$

Sönnun: Setjum $Y := (X_1 + \ldots + X_n)/n$. Þá er

$$\mathbf{E}[Y] = (\mu + \ldots + \mu)/n \quad \text{þ.e. } \mathbf{E}[Y] = \mu$$

og vegna þess að $X_1, \ldots, X_n$ eru óháðar fæst líka

$$\mathrm{Var}[Y] = (\sigma^2 + \ldots + \sigma^2)/n^2 \quad \text{þ.e. } \mathrm{Var}[Y] = \sigma^2/n\,.$$

Ójafna Chebyshevs gefur nú

$$\mathbf{P}(|Y - \mathbf{E}[Y]| > \varepsilon) \leqslant \frac{\sigma^2/n}{\varepsilon^2} \to 0, \quad n \to \infty.$$

Lögmál mikils fjölda og túlkun líkinda

Við vorum að sanna eftirfarandi reglu:

Lögmál mikils fjölda: Ef $X_1, X_2, \ldots$ eru óháðar og hafa allar sama endanlega væntigildi $\mu = \mathbf{E}[X_i]$ og allar sömu endanlegu dreifni $\sigma^2 = \mathrm{Var}[X_i]$, þá gildir fyrir öll $\varepsilon > 0$ að

$$\mathbf{P}\left(\left|\frac{X_1 + \ldots + X_n}{n} - \mu\right| > \varepsilon\right) \to 0, \quad n \to \infty.$$

Þetta skrifum við stundum losaralega svona:

$$\frac{X_1 + \ldots + X_n}{n} \approx \mu \quad \text{þegar } n \text{ er stórt.}$$

Um túlkun líkinda: Hugsum okkur að við framkvæmum tilraun í n óháð skipti og skráum hjá okkur 1 eða 0 eftir því hvort tiltekinn atburður gerist eða gerist ekki. Við fáum þá óháðar slembistærðir $X_1, \ldots, X_n$ þ.a. $\mathbf{P}(X_i = 1) = p$ þar sem p eru líkurnar á að atburðurinn gerist í einni tiltekinni tilraun (segjum þá að sú tilraun heppnist). Nú er $\mu = p$ og lögmál mikils fjölda gefur: eftir margar tilraunir er

$$\text{hlutfallsleg tíðni heppnaðra tilrauna} \approx p.$$

Þess vegna er eðlilegt að túlka líkindin p sem hlutfallslega tíðni heppnaðra tilrauna þegar tilraun er endurtekin í mörg óháð skipti.

MIKILVÆGAR STÆRÐIR

Yfirlit yfir myndskeið og efni

Skeið	lengd	Aðalsíður	hliðarsíður	efni
L22	6:24	A22		Yfirlit og Ber(p)
L23	7:01	A23		Bin(n, p)
L24	10:16	A24		Poi(λ)
L25	6:01	A25		Geo(p)
L26	7:24	A26		Hyp(N, M, n)
L27	8 02	A27		Unf [a, b]
L28	11:02	A28		Exp(λ)
L29	10:32	A29-30		Normlega stærðin
L30	6:39	A31		Varðveislureglur

Styttingar í yfirliti eru í samræmi við eftirfarandi dæmi:

L01	fyrsta myndskeið um líkindareikning
9:26	9 mínútur og 26 sekúndur
A1	fyrsta Aðalsíða
1	fyrsta hliðarsíða

Nokkrar mikilvægar slembistærðir

STRJÁLAR:	**Bernoulli-stærð**	$X \sim \mathrm{Ber}(p)$
	tvíkostastærð	$X \sim \mathrm{Bin}(n, p)$
	Poisson-stærð	$X \sim \mathrm{Poi}(\lambda)$
	strjál veldisstærð	$X \sim \mathrm{Geo}(p)$
	happadrættisstærð	$X \sim \mathrm{Hyp}(N, M, n)$
SAMFELLDAR:	**jöfn stærð**	$X \sim \mathrm{Unf}([a, b])$
	veldisstærð	$X \sim \mathrm{Exp}(\lambda)$
	normleg stærð	$X \sim \mathrm{N}(\mu, \sigma^2)$

Bernoulli-stærð $X \sim \mathrm{Ber}(p)$

Skilgreining: **Látum** $0 < p < 1$. **Slembistærð** X **kallast** ***Bernoulli-stærð*** **með stika** p **ef**

$$\mathbf{P}(X = 1) = p \quad \textbf{og} \quad \mathbf{P}(X = 0) = 1 - p.$$

Reglur: **Ef** $X \sim \mathrm{Ber}(p)$ **gildir að**

(1) $$\mathbf{E}[X] = p$$

(2) $$\mathrm{Var}[X] = p(1 - p)$$

Sönnun: (1) $\mathbf{E}[X] = 0\,\mathbf{P}(X = 0) + 1\,\mathbf{P}(X = 1) = p$.
(2) **Þar að** $0^2 = 0$ **og** $1^2 = 1$, **er** $X^2 \sim \mathrm{Ber}(p)$. **Því er** $\mathbf{E}[X^2] = \mathbf{E}[X] = p$. **Notum** $\mathrm{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2$. **Fáum** $\mathrm{Var}[X] = p - p^2 = p(1 - p)$.

Tvíkostastærð $X \sim \mathrm{Bin}(n,p)$

Skilgreining: Látum $n \geqslant 1$ og $0 < p < 1$. Slembistærð X kallast ***tvíkostastærð*** (***binomial***) með stika n, p ef

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \ldots, n.$$

Regla: Ef $X_1, \ldots, X_n$ eru **óháðar** $\mathrm{Ber}(p)$, þá er

$$X_1 + \ldots + X_n \sim \mathrm{Bin}(n,p).$$

Sönnun: Til að $\{X_1 + \ldots + X_n = k\}$ gerist þarf að fá 1 í k skipti og 0 í hin. Látum $\{i_1, \ldots, i_k\}$ vera hlutmengi úr $\{1, \ldots, n\}$ með k stökum og $\{j_1, \ldots, j_{n-k}\}$ vera fyllimengið. Notum að $X_1, \ldots, X_n$ eru óháðar $\mathrm{Ber}(p)$:

$$\mathbf{P}(X_{i_1} = 1, \ldots, X_{i_k} = 1, X_{j_1} = 0, \ldots, X_{j_{n-k}} = 0) = p^k (1-p)^{n-k}.$$

Margföldun með fjölda hlutmengja $\{i_1, \ldots, i_k\}$ gefur

$$\mathbf{P}(X_1 + \ldots + X_n = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Reglur: Ef $X \sim \mathrm{Bin}(n,p)$ gildir að

(1) $$\mathbf{E}[X] = np$$

(2) $$\mathrm{Var}[X] = np(1-p)$$

Sönnun: Látum $X_1, \ldots, X_n$ vera óháðar $\mathrm{Ber}(p)$. Notum $X_1 + \ldots + X_n \sim \mathrm{Bin}(n,p)$, summureglur um væntigildi og dreifni, og loks að $\mathbf{E}[X_i] = p$ og $\mathrm{Var}[X_i] = p(1-p)$:

$\mathbf{E}[X] = \mathbf{E}[X_1 + \ldots + X_n] = \mathbf{E}[X_1] + \ldots + \mathbf{E}[X_n] = np$,

$\mathrm{Var}[X] = \mathrm{Var}[X_1 + \ldots + X_n]$ (notum næst óhæðið)

$= \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n] = np(1-p)$.

Poisson-stærð $X \sim \mathrm{Poi}(\lambda)$

Skilgreining: Látum $0 < \lambda < \infty$. Slembistærð X kallast ***Poisson-stærð*** með stika λ ef

$$\mathbf{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

Þetta er líkindafall: $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$ svo $\sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = 1$.

Regla: Ef $Y \sim \mathrm{Bin}(n, p)$ og $X \sim \mathrm{Poi}(np)$ þá gildir að

$$|\mathbf{P}(Y \in A) - \mathbf{P}(X \in A)| \leqslant \begin{cases} np^2 \\ p \end{cases} \qquad \text{fyrir öll } A \subseteq \mathbb{R}.$$

Skrifum þetta stundum losaralega svona:

$$\mathrm{Bin}(n, p) \approx \mathrm{Poi}(np) \quad \text{þegar } p \text{ er lítið.}$$

Þetta þýðir að ef tilraun er endurtekin í n óháð skipti og líkurnar p á að heppnast í einni tiltekinni tilraun eru hverfandi litlar, þá er fjöldi heppnaðra tilrauna næstum Poisson með $\lambda = np$.

Reglur: Ef $X \sim \mathrm{Poi}(\lambda)$ gildir að

(1) $$\mathbf{E}[X] = \lambda$$

(2) $$\mathrm{Var}[X] = \lambda$$

Sönnun: (1) $\mathbf{E}[X] = \sum_0^{\infty} k\, e^{-\lambda}\frac{\lambda^k}{k!} = \lambda \sum_1^{\infty} e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!} = \lambda$.
(2) $\mathbf{E}[X(X-1)] = \sum_0^{\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} = \lambda^2 \sum_2^{\infty} e^{-\lambda}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$
gefur lokaskrefið í $\mathbf{E}[X^2] = \mathbf{E}[X(X-1)] + \mathbf{E}[X] = \lambda^2 + \lambda$
sem gefur $\mathrm{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$.

Strjál veldisstærð $X \sim \mathrm{Geo}(p)$

Skilgreining: Látum $0 < p < 1$. Slembistærð X kallast ***strjál veldisstærð*** með stika p ef

$$\mathbf{P}(X = k) = p(1-p)^{k-1}, \quad k = 1, 2, \ldots$$

Þetta er líkindafall: $\sum_{k=1}^{\infty}(1-p)^{k-1} = \frac{1}{1-(1-p)} = \frac{1}{p}$.

Regla: Látum $Y_1, Y_2, \ldots$ vera óháðar $\mathrm{Ber}(p)$. Látum X vera fyrsta n þannig að $Y_n = 1$, þ.e.

$$X := \min\{n \geqslant 1 : Y_n = 1\}.$$

Þá er $X \sim \mathrm{Geo}(p)$.

Sönnun: $\mathbf{P}(X = k) = \mathbf{P}(Y_1 = 0, \ldots, Y_{k-1} = 0, Y_k = 1)$
$= \mathbf{P}(Y_1 = 0) \ldots \mathbf{P}(Y_{k-1} = 0)\mathbf{P}(Y_k = 1) = (1-p)^{k-1}p, \; k \geqslant 1$

Regla: Ef X tekur eingöngu heiltölugildi þá:

$$X \sim \mathrm{Geo}(p) \iff \mathbf{P}(X > n) = (1-p)^n, \; n \geqslant 0$$

Sönnun á $\Rightarrow$: Notum $X \sim \mathrm{Geo}(p)$ úr reglunni hér að ofan: $\mathbf{P}(X > n) = \mathbf{P}(Y_1 = 0, \ldots, Y_n = 0) = (1-p)^n$.

Sönnun á $\Leftarrow$: $\mathbf{P}(X = k) = \mathbf{P}(X > k-1) - \mathbf{P}(X > k)$
$= (1-p)^{k-1} - (1-p)^k = (1-p)^{k-1}(1-(1-p)), \; k \geqslant 1$

Reglur: Ef $X \sim \mathrm{Geo}(p)$ gildir að

$$\mathbf{E}[X] = \frac{1}{p} \qquad \text{og} \qquad \mathrm{Var}[X] = \frac{1-p}{p^2}$$

Happadrættisstærð $X \sim \mathrm{Hyp}(N, M, n)$

Kúlur eru settar í kassa, N hvítar og M svartar.

Svo eru dregnar n kúlur af handahófi án skila.

Hverjar eru líkurnar á að nákvæmlega k séu hvítar?

Látum X vera fjölda hvítra kúlna sem dregnar eru.

Til að fá $\{X = k\}$ þarf k hvítar og $n - k$ svartar.

Hægt er að draga k úr N hvítum á $\binom{N}{k}$ marga vegu og $n - k$ úr M svörtum á $\binom{M}{n-k}$ marga vegu.

Margföldunarreglan gefur
að hægt er að fá $\{X = k\}$ á $\binom{N}{k}\binom{M}{n-k}$ marga vegu.

Heildarfjöldi möguleika á að draga n úr $N + M$ er $\binom{N+M}{n}$ svo líkurnar á að fá nákvæmlega k hvítar eru

$$\mathbf{P}(X = k) = \frac{\binom{N}{k}\binom{M}{n-k}}{\binom{N+M}{n}}$$

þar sem k uppfyllir $\max\{0, n - M\} \leqslant k \leqslant \min\{n, N\}$.

Stærð með þetta líkindafall kallast ***happadrættistærð***

Regla: Látum $X \sim \mathrm{Hyp}(N, M, n)$. Setjum $p = \frac{N}{N+M}$.
Þá gildir

$$\mathbf{E}[X] = np$$

$$\mathrm{Var}[X] = np(1-p)\left(1 - \frac{n-1}{N+M-1}\right)$$

Jöfn stærð $X \sim \mathrm{Unf}[a,b]$

Skilgreining: Látum $-\infty < a < b < \infty$. Slembistærð X kallast *jöfn stærð* á bilinu $[a,b]$ ef

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leqslant x \leqslant b, \\ 0, & \text{annars.} \end{cases}$$

Reglur: Ef $X \sim \mathrm{Unf}[a,b]$ gildir að

(1) $$F_X(x) = \frac{x-a}{b-a}, \quad a \leqslant x \leqslant b$$

(2) $$\mathbf{E}[X] = \frac{a+b}{2}$$

(3) $$\mathrm{Var}[X] = \frac{(b-a)^2}{12} \quad \text{og} \quad \sigma_X = \frac{b-a}{\sqrt{12}}$$

Sönnun: (1) $F_X(x) = \int_a^x \frac{1}{b-a}\,dy = \frac{x-a}{b-a}, \quad a \leqslant x \leqslant b.$

(2) $\mathbf{E}[X] = \int_a^b \frac{x}{b-a}\,dx = \left[\frac{1}{2}\frac{x^2}{b-a}\right]_a^b = \frac{1}{2}\frac{b^2-a^2}{b-a} = \frac{1}{2}\frac{(a+b)(b-a)}{b-a} = \frac{a+b}{2}.$

(3) $\mathbf{E}[X^2] = \int_a^b \frac{x^2}{b-a}\,dx = \frac{1}{3}\frac{b^3-a^3}{b-a} = \frac{1}{3}\frac{(a^2+ab+b^2)(b-a)}{b-a} = \frac{a^2+ab+b^2}{3}$

og því $\mathrm{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2$

$$= \frac{4(a^2+ab+b^2)-3(a^2+2ab+b^2)}{12} = \frac{a^2-2ab+b^2}{12} = \frac{(b-a)^2}{12}.$$

Varðveisluregla: $X \sim \mathrm{Unf}[a,b] \quad \Leftrightarrow \quad \dfrac{X-a}{b-a} \sim \mathrm{Unf}[0,1]$

Sönnun: Tökum $u \in [0,1]$ og $x \in [a,b]$ þ.a. $u = \frac{x-a}{b-a}$. Þá er $\mathbf{P}(X \leqslant x) = \frac{x-a}{b-a} \Leftrightarrow \mathbf{P}(\frac{X-a}{b-a} \leqslant u) = u$.

Veldisstærð $X \sim \mathrm{Exp}(\lambda)$

Skilgreining: Látum $0 < \lambda < \infty$. Slembistærð X kallast *veldisstærð* með stika λ ef

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

Reglur: Ef $X \sim \mathrm{Exp}(\lambda)$ gildir að

(1) $$\mathbf{P}(X > x) = e^{-\lambda x}, \quad x \geqslant 0$$

(2) $$F_X(x) = 1 - e^{-\lambda x}, \quad x \geqslant 0$$

(3) $$\mathbf{E}[X] = \frac{1}{\lambda}$$

(4) $$\mathrm{Var}[X] = \frac{1}{\lambda^2}$$

Sönnun: (1) $\mathbf{P}(X > x) = \int_x^\infty \lambda e^{-\lambda y} \mathbf{d}y = e^{-\lambda x}, \quad x \geqslant 0$

(2) $F_X(x) = \mathbf{P}(X \leqslant x) = 1 - \mathbf{P}(X > x) = 1 - e^{-\lambda x}, \; x \geqslant 0$

(3) $\mathbf{E}[X] = \int_0^\infty x \lambda e^{-\lambda x} \mathbf{d}x = [-x e^{-\lambda x}]_0^\infty + \int_0^\infty e^{-\lambda x} \mathbf{d}x = \frac{1}{\lambda}$

(4) $\mathbf{E}[X^2] = \int_0^\infty x^2 \lambda e^{-\lambda x} \mathbf{d}x = [-x^2 e^{-\lambda x}]_0^\infty + \int_0^\infty 2x e^{-\lambda x} \mathbf{d}x$
$= \frac{2}{\lambda} \mathbf{E}[X] = \frac{2}{\lambda^2}$ og því $\mathrm{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2}$

Skilgreining: Slembistærð X sem tekur bara jákvæð gildi kallast *minnislaus* ef

$$\mathbf{P}(X > a + b \,|\, X > a) = \mathbf{P}(X > b), \quad a, b \geqslant 0.$$

Takið eftir að ef tæki er alltaf eins og nýtt meðan það er nothæft, þá er endingartími þess minnislaus.

Regla: X er minnislaus $\Leftrightarrow$ til er λ þ.a. $X \sim \mathrm{Exp}(\lambda)$

Normleg stærð $X \sim \mathrm{N}(\mu, \sigma^2)$

Skilgreining: Látum $\mu \in \mathbb{R}$ og $0 < \sigma < \infty$. Slembistærð X kallast ***normleg stærð*** með stika μ og σ^2 ef

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad x \in \mathbb{R}.$$

Regla: $X \sim \mathrm{N}(\mu, \sigma^2) \quad \Rightarrow \quad \mathbf{E}[X] = \mu$ **og** $\mathrm{Var}[X] = \sigma^2$

Varðveisluregla: $X \sim \mathrm{N}(\mu, \sigma^2) \quad \Leftrightarrow \quad \dfrac{X-\mu}{\sigma} \sim \mathrm{N}(0,1)$

Z kallast ***stöðluð*** normleg stærð ef

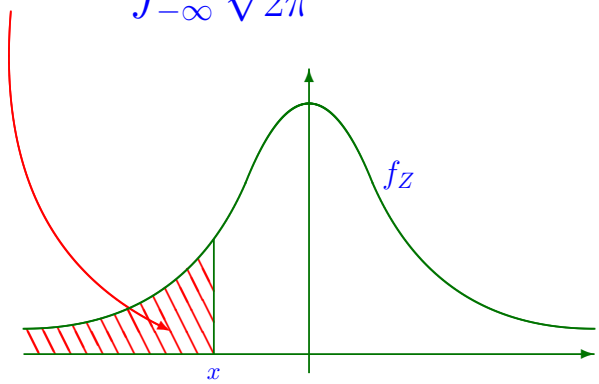
$$Z \sim \mathrm{N}(0,1).$$

Þá er $\mathbf{E}[Z] = 0$ og $\mathrm{Var}[Z] = 1$ og

$$f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Loks er $F_Z = \Phi$ þar sem

$$\Phi(x) := \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \, \mathrm{d}y, \quad x \in \mathbb{R}.$$



Notkun á $Z \sim \mathrm{N}(0,1)$ og Φ

Regla: $X \sim \mathrm{N}(\mu, \sigma^2) \;\Leftrightarrow\; F_X(x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right),\ x \in \mathbb{R}$

Normleg stærð X er samfelld og því gildir fyrir $a \in \mathbb{R}$ að $\mathbf{P}(X = a) = 0$ og $\mathbf{P}(X < a) = \mathbf{P}(X \leqslant a)$. Þess vegna má skipta á $<$ og $\leqslant$ (og á $>$ og $\geqslant$) alls staðar hér fyrir neðan.

Nota má Φ til að reikna líkur fyrir $X \sim \mathrm{N}(\mu, \sigma^2)$:

$$\mathbf{P}(X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

og

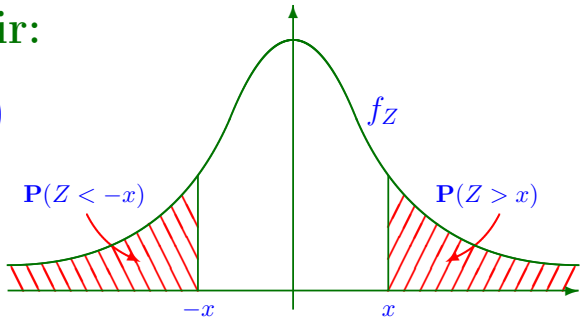
$$\mathbf{P}(a < X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Vegna samhverfu gildir:

$\mathbf{P}(Z < -x) = \mathbf{P}(Z > x)$

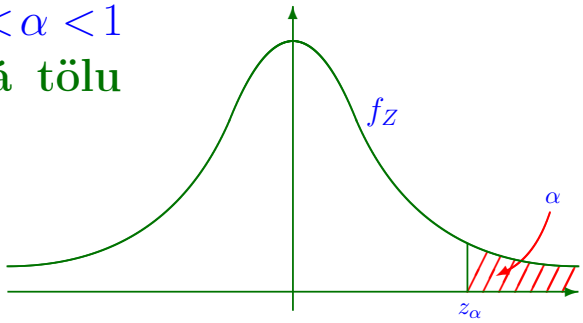
og því

$\Phi(-x) = 1 - \Phi(x)$.



Skilgreining: Fyrir $0 < \alpha < 1$ látum við z_α vera þá tölu sem uppfyllir

$$\mathbf{P}(Z > z_\alpha) = \alpha.$$



$z_{0{,}05} = 1{,}645 \quad z_{0{,}025} = 1{,}960 \quad z_{0{,}01} = 2{,}326 \quad z_{0{,}005} = 2{,}576$

Varðveislureglur

Eftirfarandi reglur gilda um óháðar stærðir X og Y.

$$\left.\begin{array}{l} X \sim \mathrm{Bin}(n,p) \\ Y \sim \mathrm{Bin}(m,p) \end{array}\right\} \Rightarrow X+Y \sim \mathrm{Bin}(n+m,p).$$

$$\left.\begin{array}{l} X \sim \mathrm{Poi}(\lambda) \\ Y \sim \mathrm{Poi}(\mu) \end{array}\right\} \Rightarrow X+Y \sim \mathrm{Poi}(\lambda+\mu).$$

$$\left.\begin{array}{l} X \sim \mathrm{N}(\mu_1,\sigma_1^2) \\ Y \sim \mathrm{N}(\mu_2,\sigma_2^2) \end{array}\right\} \Rightarrow X+Y \sim \mathrm{N}(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2).$$

$$\left.\begin{array}{l} X \sim \mathrm{Exp}(\lambda_1) \\ Y \sim \mathrm{Exp}(\lambda_2) \end{array}\right\} \Rightarrow \min\{X,Y\} \sim \mathrm{Exp}(\lambda_1+\lambda_2).$$

Sönnun á síðustu reglunni: Fyrir $x > 0$ fæst

$$\begin{aligned} \mathbf{P}(\min\{X,Y\} > x) &= \mathbf{P}(X > x, Y > x) \\ &= \mathbf{P}(X > x)\mathbf{P}(Y > x) \quad (X \textbf{ og } Y \textbf{ eru óháðar}) \\ &= e^{-\lambda_1 x} e^{-\lambda_2 x} \qquad (X \sim \mathrm{Exp}(\lambda_1) \textbf{ og } Y \sim \mathrm{Exp}(\lambda_2)) \end{aligned}$$

það er

$$\mathbf{P}(\min\{X,Y\} > x) = e^{-(\lambda_1+\lambda_2)x}, \quad x > 0.$$

Þetta þýðir að $\min\{X,Y\} \sim \mathrm{Exp}(\lambda_1+\lambda_2)$.

HÖFUÐMARKGILDISREGLAN
og
POISSONFERLI

Yfirlit yfir myndskeið og efni

Skeið	lengd	Aðalsíður	hliðarsíður	efni
L31	11:45	A32-33	33	Höfuðmarkgildisreglan og Dæmi
L32	2:08	A34		Normleg nálgun á Bin og Poi
L33	18:14	A35	35	Poissonferli og Dæmi
L34	11:06	A36		Samruni og Sundrun
L35	16:44		35-36	Dæmi – Poissonferli (A35-36)

Styttingar í yfirliti eru í samræmi við eftirfarandi dæmi:

L01	fyrsta myndskeið um líkindareikning
9:26	9 mínútur og 26 sekúndur
A1	fyrsta Aðalsíða
1	fyrsta hliðarsíða

Athugasemd um stöðlun

Ef X er slembistærð með endanlegt væntigildi μ og endanlega strangt jákvæða dreifni σ^2, kallast

$$\frac{X-\mu}{\sigma} \quad \textit{stöðluð} \text{ stærð.}$$

Þá gildir

$$\mathbf{E}\left[\frac{X-\mu}{\sigma}\right] = 0 \quad \text{og} \quad \mathrm{Var}\left[\frac{X-\mu}{\sigma}\right] = 1.$$

Látum $X_1, X_2, \ldots$ vera óháðar slembistærðir sem allar hafa sama endanlega væntigildi

$$\mu = \mathbf{E}[X_i], \quad i = 1, 2, \ldots,$$

og sömu endanlegu strangt jákvæðu dreifni

$$\sigma^2 = \mathrm{Var}[X_i], \quad i = 1, 2, \ldots.$$

Þá gildir fyrir öll $n \in \mathbb{N}$ að

$$\mathbf{E}[X_1 + \ldots + X_n] = n\mu \text{ og } \mathrm{Var}[X_1 + \ldots + X_n] = n\sigma^2$$

og því gefur stöðlun

$$\mathbf{E}\left[\frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}\right] = 0$$

og

$$\mathrm{Var}\left[\frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}\right] = 1.$$

Höfuðmarkgildisreglan – Normleg nálgun

Höfuðmarkgildisreglan: Látum $X_1, X_2, \ldots$ vera óháðar slembistærðir sem allar hafa sama dreififall með endanlegt væntigildi

$$\mu = \mathbf{E}[X_i], \quad i = 1, 2, \ldots,$$

og endanlega strangt jákvæða dreifni

$$\sigma^2 = \mathrm{Var}[X_i], \quad i = 1, 2, \ldots.$$

Þá gildir fyrir öll $x \in \mathbb{R}$ að

$$\mathbf{P}\left(\frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}} \leqslant x\right) \to \Phi(x), \quad n \to \infty.$$

Við skrifum þetta oft losaralega svona:

$$\frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}} \approx \mathrm{N}(0,1) \quad \text{þegar } n \text{ stórt}$$

eða svona:

$$X_1 + \ldots + X_n \approx \mathrm{N}(n\mu, n\sigma^2) \quad \text{þegar } n \text{ stórt.}$$

Dæmi: Byggja á 1000 íbúða hverfi. Vitað er að bílafjöldi á dæmigerða fjölskyldu í svona hverfi hefur eftirfarandi dreifingu:

20% á engan bíl,
70% á einn bíl,
10% á tvo bíla.

Spurning: Hvað þarf stóra bílageymslu til að hún dugi fyrir hverfið með u.þ.b. 95% líkum?

Normleg nálgun á $\mathrm{Bin}(n,p)$ og á $\mathrm{Poi}(\lambda)$

Höfuðmarkgildisreglan segir að ef $X_1, \ldots, X_n$ eru óháðar slembistærðir sem allar hafa sama dreififall með endanlegt væntigildi μ og endanlega strangt jákvæða dreifni σ^2 þá gildir að

$$X_1 + \ldots + X_n \approx \mathrm{N}(n\mu, n\sigma^2) \qquad \text{þegar } n \text{ stórt.}$$

Vegna þess að $\mathrm{Bin}(n,p)$ er summa n óháðra $\mathrm{Ber}(p)$ leiðir af þessu að

$$\mathrm{Bin}(n,p) \approx \mathrm{N}(np, np(1-p)) \quad \text{þegar } n \text{ er stórt.}$$

Hversu stórt n þarf að vera fer eftir því hvað p er. Nálgunin verður þeim mun betri sem p er nær $\frac{1}{2}$. Hér er algeng puttaregla:

$$\boxed{\text{puttaregla:} \quad np(1-p) \geqslant 10}$$

Vegna þess að $\mathrm{Poi}(\lambda) \approx \mathrm{Bin}(n, \lambda/n)$ þegar n er stórt leiðir af þessu að

$$\mathrm{Poi}(\lambda) \approx \mathrm{N}(\lambda, \lambda) \quad \text{þegar } \lambda \text{ er stórt.}$$

Hér er algeng puttaregla:

$$\boxed{\text{puttaregla:} \ \lambda \geqslant 10}$$

Poisson-ferli (c)

Látum $T_1, T_2, \ldots$ vera slembistærðir þ.a.

$$0 < T_1 < T_2 < \ldots \quad \text{og} \quad T_n \to \infty, \ n \to \infty.$$

Hugsum okkur að þessi T_n séu tímapunktar þegar einhver viðburður gerist.

Talningarferli viðburðastreymisins er $(N_t)_{t \geqslant 0}$ þar sem

$N_t = \#\{n \geqslant 1 : T_n \leqslant t\} =$ **fjöldi viðburða á bilinu** $[0, t]$

Billengdir viðburðastreymisins eru

$$X_1 = T_1 \quad \text{og} \quad X_n = T_n - T_{n-1}, \ n \geqslant 2.$$

Skilgreining: Látum $0 < c < \infty$. Talningarferli $(N_t)_{t \geqslant 0}$ kallast ***Poisson-ferli*** með ***tíðni*** c ef

- $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}$ eru óháðar slembistærðir fyrir öll $n \geqslant 1$ og $0 < t_1 < t_2 < \ldots < t_n$.
- $\mathbf{E}[N_t] = ct$ fyrir öll $t \geqslant 0$.

Regla: Ef $(N_t)_{t \geqslant 0}$ er Poisson-ferli með tíðni c gildir:

$$N_t \sim \mathrm{Poi}(ct), \quad t \geqslant 0.$$

Jafngildisregla: $(N_t)_{t \geqslant 0}$ er Poisson-ferli með tíðni c

$$\Leftrightarrow$$

$X_1, X_2, \ldots$ eru óháðar og allar $\mathrm{Exp}(c)$

$$\Leftrightarrow$$

$N_t \sim \mathrm{Poi}(ct)$, $t \geqslant 0$, og ef skilyrt er með $\{N_t = n\}$ er eins og þessum n punktum hafi verið stráð af handahófi á bilið $[0, t]$.

Poisson-ferli – samruni og sundrun

Regla (samruni): Látum $(L_t)_{t\geqslant 0}$ og $(M_t)_{t\geqslant 0}$ vera óháð Poisson-ferli með tíðnir a og b. Setjum

$$N_t = L_t + M_t, \quad t \geqslant 0,$$

og fyrir $n \geqslant 1$

$$I_n = \begin{cases} 0 & \text{ef } n^{\text{ti}} \text{ punktur } (N_t)_{t\geqslant 0} \text{ kemur úr } (L_t)_{t\geqslant 0}, \\ 1 & \text{ef } n^{\text{ti}} \text{ punktur } (N_t)_{t\geqslant 0} \text{ kemur úr } (M_t)_{t\geqslant 0}. \end{cases}$$

Þá gildir að

- $(N_t)_{t\geqslant 0}$ er Poisson-ferli með tíðni $c = a + b$,
- $I_1, I_2, \ldots$ eru óháðar og allar $\mathrm{Ber}(p)$ þar sem $p = \frac{b}{c}$,
- og runan $(I_1, I_2, \ldots)$ er óháð $(N_t)_{t\geqslant 0}$.

Regla (sundrun): Látum

- $(N_t)_{t\geqslant 0}$ vera Poisson-ferli með tíðni c,
- $I_1, I_2, \ldots$ vera óháðar og allar $\mathrm{Ber}(p)$,
- og rununa $(I_1, I_2, \ldots)$ vera óháða $(N_t)_{t\geqslant 0}$.

Látum $(L_t)_{t\geqslant 0}$ og $(M_t)_{t\geqslant 0}$ vera talningarferli þannig að

$$L_t + M_t = N_t, \quad t \geqslant 0,$$

og

$$I_n = \begin{cases} 0 & \text{ef } n^{\text{ti}} \text{ punktur } (N_t)_{t\geqslant 0} \text{ kemur úr } (L_t)_{t\geqslant 0}, \\ 1 & \text{ef } n^{\text{ti}} \text{ punktur } (N_t)_{t\geqslant 0} \text{ kemur úr } (M_t)_{t\geqslant 0}. \end{cases}$$

Þá eru $(L_t)_{t\geqslant 0}$ og $(M_t)_{t\geqslant 0}$ óháð Poisson-ferli með tíðnir $a = (1-p)c$ og $b = pc$.

---

hliðarsíður

Yfirlit yfir hliðarsíður: myndskeið og efni

Skeið	lengd	Aðalsíður	**hliðarsíður**	efni
L01	9:26		**1-2**	Inngangur um líkinda- og tölfræði
L04	11:03		**3-4**	Talningarfræði – tengist A3
L05	2:45		**5**	Dæmi – Jafnar líkur (brids, A3)
L06	14:19	A4	**7**	Dæmi – Skilyrtar líkur
L14	17:17		**9-11**	Dæmi – tengist A11-13
L15	9:14	A14	**13-15**	Væntigildi og miðgildi
L17	13:28	A16	**17-22**	Sannanir á reglum um væntigildi
L19	13:13	A17	**23**	Sannanir á reglum um dreifni
L20	15:01	A18	**25-31**	Sannanir á reglum um samdreifni
L31	11:45	A33	**33**	Dæmi – Höfuðmarkgildisreglan
L35	16:44		**35-36**	Dæmi – Poissonferli (A35-36)

Einstaka síður með jöfnum númerum eru auðar

Styttingar í yfirliti eru í samræmi við eftirfarandi dæmi:
L01 fyrsta myndskeið um líkindareikning
9:26 9 mínútur og 26 sekúndur
A1 fyrsta Aðalsíða
1 fyrsta hliðarsíða

Þessi opna er inngangur að Aðalsíðum – L01

Líkindafræði – Lögmál tilviljunarinnar

Allt er tilviljun háð, eða flest.

En jafnvel tilviljunin lýtur lögmálum.

Líkindafræðin er sú grein stærðfræðinnar sem fjallar um þessi lögmál á sama hátt og til dæmis rúmfræðin fjallar um lögmál flatarmynda og rúmforma.

Tölfræði – Listin að lesa úr gögnum

Tölfræðin byggir á líkindafræðinni
og er einskonar andhverfa hennar:

Líkindafræðin spáir fyrir um
óþekktar tilraunaniðurstöður
á grundvelli
þekktra aðstæðna.

Tölfræðin snýr þessu við, hún metur
óþekktar aðstæður
á grundvelli
þekktra tilraunaniðurstaðna (gagna).

Dæmi um samband líkindafræði og tölfræði

Hugsum okkur að gerð sé skoðanakönnun þar sem 1000 handahófsvaldir kjósendur eru spurðir hvort þeir ætli að kjósa tiltekinn flokk eða ekki. Setjum

$p =$ hlutfallsfylgi flokksins meðal kjósenda

og

$\hat{p} =$ hlutfallsfylgi flokksins meðal spurðra.

Líkindafræðin spáir fyrir um $\hat{p}$ á grundvelli p,

(*) $$\hat{p} \approx p.$$

Tölfræðin snýr þessu við og metur p á grundvelli $\hat{p}$,

$$p \approx \hat{p}.$$

Líkindafræðin segir einnig til um nákvæmnina: T.d. ef margir óháðir aðilar gerðu svona könnun þá væri

$$\hat{p} = p \pm 1,96\sqrt{\frac{p(1-p)}{1000}} \quad \text{í um það bil 95\% skipta.}$$

Tölfræðin snýr þessu líka við:

$$p = \hat{p} \pm 1,96\sqrt{\frac{\hat{p}(1-\hat{p})}{1000}} \quad \text{í um það bil 95\% skipta,}$$

þar sem sett var inn $p(1-p) \approx \hat{p}(1-\hat{p})$ með leyfi (*).

Ef í eitthvert skiptið sem svona skoðanakönnun er gerð fæst t.d. $\hat{p} = 20\%$ þá gefur síðasta formúlan að

$$p = 0,2 \pm 1,96\sqrt{\frac{0,2(1-0,2)}{1000}} \quad \text{með u.þ.b. 95\% öryggi}$$

þ.e. með u.þ.b. 95% öryggi er fylgið $17,5\%$ til $22,5\%$.

Þessi opna kemur á eftir A3 um jöfnar líkur – L04

Úr talningarfræði

Þegar reikna á $\mathbf{P}(A) = \frac{\text{fjöldinn í } A}{\text{fjöldinn í } \Omega}$ þá þarf að telja stökin bæði í A og Ω. Þessi mengi geta verið mjög stór (svo ekki sé meira sagt) og flókin. Lykillinn er:

Grundvallarregla talningarfræðinnar: Ef gera má eina aðgerð á l vegu og svo aðra aðgerð á m vegu þá má gera heildaraðgerðina á $l \cdot m$ vegu.

Skilgreinum n **hrópmerkt** svona: $0! = 1$ og

$$n! = 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n \quad \textbf{fyrir} \quad n \geqslant 1$$

Skilgreinum n **yfir** k svona:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad \textbf{fyrir} \quad 0 \leqslant k \leqslant n$$

Á hve marga vegu má velja k stök úr n stökum?

	með skilum	**án skila**
með tilliti til raðar	(1) n^k	(2) $\frac{n!}{(n-k)!}$
án tillits til raðar	(3) $\binom{n+k-1}{k}$	(4) $\binom{n}{k}$

Sönnun á (1): Þurfum að telja hve margar runur má mynda með því að velja eitt stak af n í fyrsta sæti, eitt stak af n í annað sæti, . . . , og að lokum eitt stak af n í sæti númer k. Samkvæmt Grundvallarreglunni er fjöldinn n margfaldað saman k sinnum.

Sönnun á (2): Þurfum að telja hve margar runur má mynda með því að velja eitt stak af n í fyrsta sæti, eitt stak af þeim $n-1$ sem þá eru eftir í annað sæti, ..., og að lokum eitt stak af þeim $n-(k-1)$ sem eru eftir þegar búið er að velja $k-1$ stök. Samkvæmt Grundvallarreglunni er fjöldinn $n(n-1)\ldots(n-(k-1))$ þ.e.a.s. $n(n-1)\ldots(n-k+1)$. Margföldun með $(n-k)!/(n-k)! = 1$ gefur (2).

Tilvik (4) má umorða: **Á hve marga vegu má velja hlutmengi með k stökum úr mengi með n stökum?**

Sönnun á (4): Köllum þennan möguleikafjölda x. Ef við veljum fyrst hlutmengi með k stökum (sem má gera á x vegu) og röðum svo stökunum (sem má gera á $k!$ vegu) þá erum við komin með tilvik (2). Grundvallarreglan gefur $x \cdot k! = \frac{n!}{(n-k)!}$ þ.e. $x = \binom{n}{k}$.

Sönnun á (3): Ef við lítum á stökin n sem skálar og stökin k sem kúlur, þá getum við umorðað (3) svona: Á hve marga vegu má leggja k eins kúlur í n mismunandi skálar? (Athugið að hér geta allt að k kúlur farið í eina skál). Þessa spurningu má svo aftur umorða: Á hve marga vegu má setja $n-1$ skilrúm á milli k kúlna? Ef við lítum á kúlur og skilrúm sem $k+n-1$ staði þá þurfum við að velja k staði (án skila og án tillilts til raðar) til að setja kúlur á (og á hina staðina $n-1$ setjum við skilrúm). Þetta má einmitt gera á $\binom{k+n-1}{k}$ vegu.

Hér er talningarreglum beitt á jafnar líkur – L05

Dæmi: Í brids fær spilamaður 13 handahófsvalin spil úr venjulegum 52 spila stokki. Hverjar eru líkurnar á að hann fái nákvæmlega 7 spaða?

Lausn: Setjum

$\Omega =$ safn allra 13-spila hlutmengja úr stokknum

og

$A =$ safn allra 13-spila hlutmengja úr stokknum með 7 spöðum og 6 ekki-spöðum.

Látum $\mathbf{P}$ vera jafnar líkur. Þá fæst

$$\mathbf{P}(A) = \frac{\text{fjöldinn í } A}{\text{fjöldinn í } \Omega} = \frac{\binom{13}{7}\binom{39}{6}}{\binom{52}{13}} = \ldots = 0,0089 < 1\%$$

Athugasemd: Sama gildir fyrir hjarta, tígul og lauf. Vegna þess að þessir 4 atburðir eru sundurlægir fæst að líkurnar á að fá sjölit eru um það bil $3\frac{1}{2}\%$.

Þessi síða sýnir notkun á skilyrðingu – A4 – L06

Dæmi: Vitað er að í ákveðnum þjóðfélgashópi eru 0,4% með tiltekið krabbamein. Gerð er forkönnun á öllum í hópnum með aðferð sem gefur rétta niðurstöðu með 95% líkum hvort sem manneskjan sem verið er að skoða er með krabbann eða ekki.

Hverjar eru líkurnar á að manneskja, sem greinist með krabbann, sé í rauninni með hann?

Lausn: Setjum

A = manneskjan greinist með krabbann

B = manneskjan er með krabbann

Spurt er um skilyrtu líkurnar $\mathbf{P}(B|A)$.
Við vitum að

$$\mathbf{P}(B) = 0,004 \quad \Longrightarrow \quad \mathbf{P}(B^c) = 0,996$$
$$\mathbf{P}(A|B) = 0,95$$
$$\mathbf{P}(A^c|B^c) = 0,95 \quad \Longrightarrow \quad \mathbf{P}(A|B^c) = 0,05$$

Notum nú Bayes-lögmálið og svo Skiptingarregluna í fyrstu tvö skrefin í eftirfarandi reikingum:

$$\mathbf{P}(B|A) = \frac{\mathbf{P}(B)\mathbf{P}(A|B)}{\mathbf{P}(A)} = \frac{\mathbf{P}(B)\mathbf{P}(A|B)}{\mathbf{P}(B)\mathbf{P}(A|B) + \mathbf{P}(B^c)\mathbf{P}(A|B^c)}$$
$$= \frac{0,004 \cdot 0,95}{0,004 \cdot 0,95 + 0,996 \cdot 0,05} = 0,07 = 7\%$$

Athugasemd: Sama á við um gæðakannanir þar sem leitað er að gölluðum vörum. Í því tilviki er

A = varan greinist gölluð

B = varan er gölluð

Næstu tvær opnur: samþéttleiki – A11-13 – L14

Dæmi: Slembistærðirnar X og Y hafa samþéttleikann

$$f(x,y) = \begin{cases} 2, & 0 \leqslant x < y \leqslant 1 \\ 0, & \text{annars.} \end{cases}$$

Sýnið að (a) f er þéttleiki og ákvarðið (b) f_X og (c) f_Y.

Lausn: (a) $f \geqslant 0$ og $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\mathrm{d}x\mathrm{d}y = \int_0^1\int_0^y 2\mathrm{d}x\mathrm{d}y = 1$

(b) $f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y = \int_x^1 2\,\mathrm{d}y = 2(1-x), \quad 0 \leqslant x \leqslant 1$

(c) $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x = \int_0^y 2\,\mathrm{d}x = 2y, \quad 0 \leqslant y \leqslant 1$

Athugasemd 1: Stundum er látið nægja að skrifa t.d. $f_Y(y) = 2y$, $0 \leqslant y \leqslant 1$, þegar $f_Y(y) = 0$ annars. Í (c) er ljóst að $f_Y(y) = 0$ ef $y < 0$ eða $y > 1$.

Athugasemd 2: Ef C er hlutmengi í planinu $\mathbb{R}^2$ með flatarmál $0 < c < \infty$ og samþéttleiki X og Y er

$$f(x,y) = \begin{cases} 1/c, & (x,y) \in C, \\ 0, & \text{annars,} \end{cases}$$

þá er sagt að (X,Y) hafi ***jafna dreifingu*** á C og líka að punkturinn (X,Y) hafi verið valinn af ***handahófi*** úr C. Til dæmis er (X,Y) hér að ofan með jafna dreifingu á þríhyrningnum $C = \{(x,y) : 0 \leqslant x < y \leqslant 1\}$.

Dæmi (framhald): Slembistærðirnar X og Y hafa samþéttleikann

$$f(x,y) = \begin{cases} 2, & 0 \leqslant x < y \leqslant 1 \\ 0, & \textbf{annars.} \end{cases}$$

Við leiddum út áðan að

$$f_X(x) = 2(1-x), \quad 0 \leqslant x \leqslant 1,$$

$$f_Y(y) = 2y, \qquad 0 \leqslant y \leqslant 1.$$

Eru (d) X og Y óháðar?

Lausn: (d) Ákvörðum fyrst $f_X f_Y$,

$$f_X(x) f_Y(y) = \begin{cases} 4(1-x)y, & 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1 \\ 0, & \textbf{annars} \end{cases}$$

Þetta er greinilega ekki samþéttleiki X og Y: til dæmis setur $f_X f_Y$ líkur sem eru minni en 1 á þríhyrninginn $C = \{(x,y) : 0 \leqslant x < y \leqslant 1\}$,

$$\int_0^1 \int_0^y 4(1-x)y \mathrm{d}x \mathrm{d}y < 1$$

en f setur líkurnar 1 á C,

$$\mathbf{P}((X,Y) \in C) = \int_0^1 \int_0^y 2 \mathrm{d}x \mathrm{d}y = 1.$$

Því eru X og Y ekki óháðar.

Dæmi (meira framhald): Slembistærðirnar X og Y hafa samþéttleikann

$$f(x,y) = \begin{cases} 2, & 0 \leqslant x < y \leqslant 1 \\ 0, & \text{annars.} \end{cases}$$

Við leiddum út áðan að X og Y eru háðar og að

$$f_X(x) = 2(1-x), \quad 0 \leqslant x \leqslant 1, \qquad (= 0 \text{ annars})$$

$$f_Y(y) = 2y, \qquad 0 \leqslant y \leqslant 1, \qquad (= 0 \text{ annars}).$$

Ákvarðið (e) $f_{X|Y}(x|y)$ og (f) $f_{Y|X}(y|x)$.

Lausn: (e) Tökum fyrst eftir að $f_Y(y) > 0$ ef og aðeins ef $0 < y \leqslant 1$. Fyrir þessi gildi á y fæst nú

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} 1/y, & 0 \leqslant x < y, \\ 0, & \text{annars.} \end{cases}$$

(f) Tökum fyrst eftir að $f_X(x) > 0$ ef og aðeins ef $0 \leqslant x < 1$. Fyrir þessi gildi á x fæst nú

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} 1/(1-x), & x \leqslant y \leqslant 1, \\ 0, & \text{annars.} \end{cases}$$

Miðgildi m

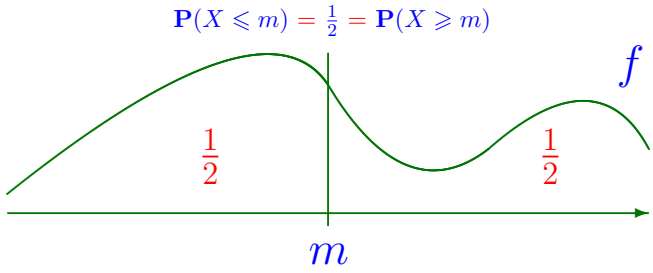
Hvernig á að lýsa staðsetningu slembistærðar X með einni tölu?

Skilgreining: Tala m sem er þannig að

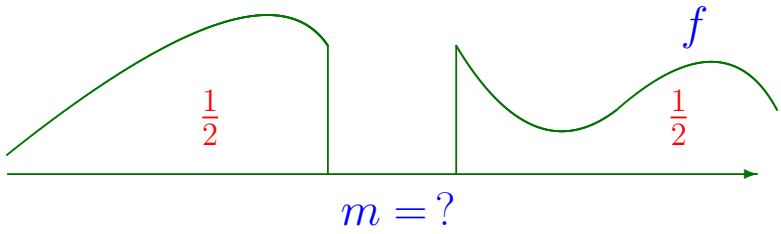
$$\mathbf{P}(X \leqslant m) \geqslant \tfrac{1}{2} \leqslant \mathbf{P}(X \geqslant m)$$

kallast miðgildi X. Skrifum stundum m_X í stað m.

Dæmi þegar X er með þéttleika f:



Miðgildi er ekki alltaf heppilegur mælikvarði á staðsetningu. Til dæmis geta verið til mörg miðgildi:



Athugasemd um rithátt fyrir summur.

Ef g er raungilt fall og $g(x) = 0$ þegar $x \notin \{a_1, \ldots, a_n\}$ þá skrifum við

$$\sum_x g(x) := \sum_{k=1}^{n} g(a_k)$$

Og ef $g(x) = 0$ þegar $x \notin \{a_1, a_2 \ldots\}$ þá skrifum við

$$\sum_x g(x) := \sum_{k=1}^{\infty} g(a_k)$$

svo fremi $\sum_{k=1}^{\infty} g(a_k)$ sé til (þ.e. röðin sé samleitin).

Dæmi: Ef X er strjál slembistærð sem tekur gildi í $\{a_1, \ldots, a_n\}$ þá er *væntigildi* X skilgreint svona

$$\sum_x x\mathbf{P}(X = x) := \sum_{k=1}^{n} a_k \mathbf{P}(X = a_k)$$

Ef X er strjál slembistærð sem tekur gildi í $\{a_1, a_2, \ldots\}$ þá er *væntigildi* X skilgreint svona

$$\sum_x x\mathbf{P}(X = x) := \sum_{k=1}^{\infty} a_k \mathbf{P}(X = a_k)$$

Miðgildi m

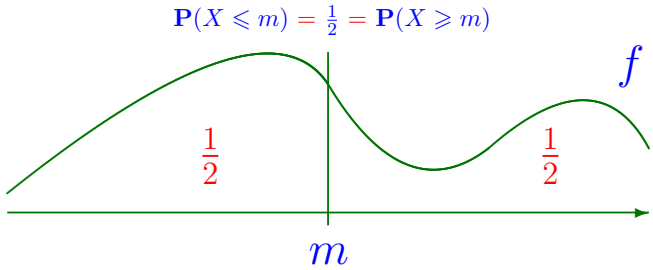
Hvernig á að lýsa staðsetningu slembistærðar X með einni tölu?

Skilgreining: Tala m sem er þannig að

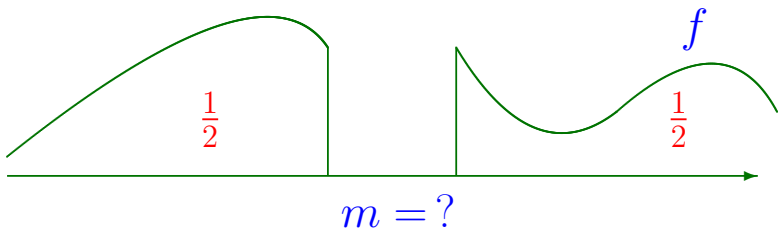
$$\mathbf{P}(X \leqslant m) \geqslant \tfrac{1}{2} \leqslant \mathbf{P}(X \geqslant m)$$

kallast miðgildi X. Skrifum stundum m_X í stað m.

Dæmi þegar X er með þéttleika f:



Miðgildi er ekki alltaf heppilegur mælikvarði á staðsetningu. Til dæmis geta verið til mörg miðgildi:



Dæmi um miðgildi m og væntigildi μ.

Skoðum þéttleikann $f(x) = e^{-x},\ x \geqslant 0$:

$$\int_m^\infty e^{-x}\mathbf{d}x = [-e^{-x}]_m^\infty = e^{-m} = \tfrac{1}{2} \qquad \Rightarrow m = \ln 2 \approx 0,7$$

$$\int_0^\infty xe^{-x}\mathbf{d}x = [-xe^{-x}]_0^\infty + \int_0^\infty e^{-x}\mathbf{d}x = 1 \qquad \Rightarrow \mu = 1$$

Næstu þrjár opnur: væntigildisreglur – A16 – L17

Lögmál ómeðvitaða tölfræðingsins, strjála tilvikið:
Ef $X_1, \ldots, X_n$ eru strjálar og $Y = g(X_1, \ldots, X_n)$ þá gildir

$$\mathbf{E}[Y] = \sum_{x_n} \cdots \sum_{x_1} g(x_1, \ldots, x_n) \mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

Sönnun: Skilgreinum fall δ_a svona

$$\delta_a(y) = \begin{cases} 1 & \text{ef } y = a, \\ 0 & \text{ef } y \neq a. \end{cases}$$

Takið eftir að

$$\mathbf{P}(Y = y) = \sum_{x_n} \cdots \sum_{x_1} \delta_{g(x_1,\ldots,x_n)}(y) \mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

Þetta gefur annað skrefið í eftirfarandi reikningi

$$\mathbf{E}[Y] = \sum_y y \mathbf{P}(Y = y) \quad \text{skv. skilgreiningu á væntigildi}$$

$$= \sum_y y \left(\sum_{x_n} \cdots \sum_{x_1} \delta_{g(x_1,\ldots,x_n)}(y) \mathbf{P}(X_1 = x_1, \ldots, X_n = x_n) \right)$$

$$= \sum_{x_n} \cdots \sum_{x_1} \left(\sum_y y \delta_{g(x_1,\ldots,x_n)}(y) \right) \mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

Takið loks eftir að

$$\sum_y y \delta_{g(x_1,\ldots,x_n)}(y) = g(x_1, \ldots, x_n)$$

vegna þess að allir liðirnir í rauðu summunni eru 0 nema liðurinn þegar $y = g(x_1, \ldots, x_n)$.

Skoðum strjál X (samskonar gildir fyrir samfelld X):

Skilgreining á væntigildi: $\mathbf{E}[X] := \sum_x x\mathbf{P}(X = x)$

Lögmál ómeðvitaða tölfræðingsins: $Y = g(X_1, \ldots, X_n)$,

$$\mathbf{E}[Y] = \sum_{x_n} \cdots \sum_{x_1} g(x_1, \ldots, x_n)\mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

Reglur: **Fyrir** $a, a_1, \ldots, a_n \in \mathbb{R}$ **gildir**

(1) $\mathbf{E}[a] = a$.

Sönnun á (1): $\mathbf{E}[a] = \mathbf{P}(a = a)a = 1a = a$

Skoðum strjál X (samskonar gildir fyrir samfelld X):

Skilgreining á væntigildi: $\mathbf{E}[X] := \sum_x x\mathbf{P}(X = x)$

Lögmál ómeðvitaða tölfræðingsins: $Y = g(X_1, \ldots, X_n)$,

$$\mathbf{E}[Y] = \sum_{x_n} \cdots \sum_{x_1} g(x_1, \ldots, x_n)\mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

Reglur: Fyrir $a, a_1, \ldots, a_n \in \mathbb{R}$ gildir

(1) $\mathbf{E}[a] = a$.

(2) $\mathbf{E}[aX] = a\mathbf{E}[X]$.

Sönnun á (2): Setjum $g(x) = ax$:

$$\mathbf{E}[aX] = \sum_x ax\mathbf{P}(X = x) = a\sum_x x\mathbf{P}(X = x) = a\mathbf{E}[X]$$

Skoðum strjál X (samskonar gildir fyrir samfelld X):

Skilgreining á væntigildi: $\mathbf{E}[X] := \sum_x x\mathbf{P}(X = x)$

Lögmál ómeðvitaða tölfræðingsins: $Y = g(X_1, \ldots, X_n)$,

$$\mathbf{E}[Y] = \sum_{x_n} \cdots \sum_{x_1} g(x_1, \ldots, x_n)\mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

Reglur: Fyrir $a, a_1, \ldots, a_n \in \mathbb{R}$ gildir

(1) $\mathbf{E}[a] = a$.

(2) $\mathbf{E}[aX] = a\mathbf{E}[X]$.

(3) $\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y]$.

Sönnun á (3): Setjum $g(x, y) = x + y$:

$$\mathbf{E}[X + Y] = \sum_x \sum_y (x + y)\mathbf{P}(X = x, Y = y)$$

$$= \sum_x \sum_y \Big(x\mathbf{P}(X = x, Y = y) + y\mathbf{P}(X = x, Y = y) \Big)$$

$$= \sum_x \sum_y x\mathbf{P}(X = x, Y = y) + \sum_x \sum_y y\mathbf{P}(X = x, Y = y)$$

$$= \sum_x x \sum_y \mathbf{P}(X = x, Y = y) + \sum_y y \sum_x \mathbf{P}(X = x, Y = y)$$

$$= \sum_x x\mathbf{P}(X = x) \qquad + \sum_y y\mathbf{P}(Y = y)$$

$$= \mathbf{E}[X] \qquad + \mathbf{E}[Y]$$

Skoðum strjál X (samskonar gildir fyrir samfelld X):

Skilgreining á væntigildi: $\mathbf{E}[X] := \sum_x x\mathbf{P}(X = x)$

Lögmál ómeðvitaða tölfræðingsins: $Y = g(X_1, \ldots, X_n)$,

$$\mathbf{E}[Y] = \sum_{x_n} \cdots \sum_{x_1} g(x_1, \ldots, x_n)\mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

Reglur: Fyrir $a, a_1, \ldots, a_n \in \mathbb{R}$ gildir

(1) $\mathbf{E}[a] = a$.

(2) $\mathbf{E}[aX] = a\mathbf{E}[X]$.

(3) $\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y]$.

(4) $\mathbf{E}[a_1X_1 + \ldots + a_nX_n] = a_1\mathbf{E}[X_1] + \ldots + a_n\mathbf{E}[X_n]$

Sönnun á (4): Notum fyrst (3) ítrekað

$$\begin{aligned}\mathbf{E}[a_1X_1 + \ldots + a_nX_n] &= \mathbf{E}[a_1X_1 + \ldots a_{n-1}X_{n-1}] + \mathbf{E}[a_nX_n] \\ &= \ldots = \mathbf{E}[a_1X_1] + \ldots + \mathbf{E}[a_nX_n]\end{aligned}$$

og notum svo (2).

Skoðum strjál X (samskonar gildir fyrir samfelld X):

Skilgreining á væntigildi: $\mathbf{E}[X] := \sum_x x\mathbf{P}(X = x)$

Lögmál ómeðvitaða tölfræðingsins: $Y = g(X_1, \ldots, X_n)$,

$$\mathbf{E}[Y] = \sum_{x_n} \cdots \sum_{x_1} g(x_1, \ldots, x_n)\mathbf{P}(X_1 = x_1, \ldots, X_n = x_n)$$

Regla: Ef X og Y eru óháð gildir $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$.

Sönnun: Setjum $g(x, y) = xy$,

$$\begin{aligned}
\mathbf{E}[XY] &= \sum_x \sum_y xy\mathbf{P}(X = x, Y = y) \\
&= \sum_x \sum_y xy\mathbf{P}(X = x)\mathbf{P}(Y = y) \quad (X \text{ og } Y \text{ óháðar}) \\
&= \sum_x x\mathbf{P}(X = x) \sum_y y\mathbf{P}(Y = y) \\
&= \mathbf{E}[X] \qquad\qquad \mathbf{E}[Y]
\end{aligned}$$

Á þessari síðu: sannanir á dreifnireglum – A17 – L19

Skilgreining: ***Dreifni*** **slembistærðar** X **er**

$$\mathrm{Var}[X] := \mathbf{E}[(X-\mu)^2] \qquad (\textbf{hér er } \mu = \mathbf{E}[X])$$

Reglur:

(1) $$\mathrm{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2.$$

(2) $$\mathrm{Var}[a+bX] = b^2\mathrm{Var}[X], \qquad a, b \in \mathbb{R}.$$

Sönnun á (1): **Notum væntigildisreglur fyrir þriðja skrefið hér á eftir og** $\mu = \mathbf{E}[X]$ **fyrir það fjórða**

$$\begin{aligned}\mathrm{Var}[X] &= \mathbf{E}[(X-\mu)^2] = \mathbf{E}[X^2 - 2\mu X + \mu^2] \\ &= \mathbf{E}[X^2] - 2\mu\mathbf{E}[X] + \mu^2 = \mathbf{E}[X^2] - (\mathbf{E}[X])^2\end{aligned}$$

Sönnun á (2): **Notum** $\mathbf{E}[a+bX] = a + b\mathbf{E}[X]$ **fyrir fyrsta skrefið og væntigildisreglu fyrir það fjórða**

$$\begin{aligned}\mathrm{Var}[a+bX] &= \mathbf{E}[\big((a+bX) - (a+b\mathbf{E}[X])\big)^2] \\ &= \mathbf{E}[(bX - b\mathbf{E}[X])^2] = \mathbf{E}[b^2(X-\mathbf{E}[X])^2] \\ &= b^2\mathbf{E}[(X-\mathbf{E}[X])^2] = b^2\mathrm{Var}[X]\end{aligned}$$

Regla: **Ef** $X_1, \ldots, X_n$ **eru óháðar gildir að**

$$\mathrm{Var}[X_1 + \ldots + X_n] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n].$$

Sönnun: **Sjá neðst á hliðarsíðu 31.**

Næstu fjórar opnur: samdreifnireglur – A18 – L20

Skilgreining: ***Samdreifni*** **slembistærða** X **og** Y **er**

$$\mathrm{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

og ***fylgnistuðull*** **þeirra er**

$$\rho = \rho_{X,Y} := \frac{\mathrm{Cov}[X,Y]}{\sigma_X \sigma_Y}\,.$$

Takið eftir að $\mathrm{Cov}[X,X] = \mathrm{Var}[X]$.

Regla: $-1 \leqslant \rho \leqslant 1$ **(Sönnum þetta ekki.)**

Regla: $\mathrm{Cov}[X,Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$

Sönnun: $\mathrm{Cov}[X,Y] = \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$

$= \mathbf{E}[XY - X\mu_Y - \mu_X Y + \mu_X\mu_Y]$

$= \mathbf{E}[XY] - \mathbf{E}[X]\mu_Y - \mu_X\mathbf{E}[Y] + \mu_X\mu_Y$ **(væntigildisreglur)**

$= \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$

Skilgreining: ***Samdreifni*** **slembistærða X og Y er**

$$\mathrm{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

Reglur: **Fyrir $a \in \mathbb{R}$ og $n, m \in \mathbb{N}$ gildir að**

$$\mathrm{Cov}[aX,Y] = a\,\mathrm{Cov}[X,Y] \tag{1}$$

Sönnun á (1): **Notum $\mu_{aX} = \mathbf{E}[aX] = a\mathbf{E}[X] = a\mu_X$,**

$$\begin{aligned}
\mathrm{Cov}[aX,Y] &= \mathbf{E}[(aX - a\mu_X)(Y-\mu_Y)] \\
&= \mathbf{E}[a(X-\mu_X)(Y-\mu_Y)] \\
&= a\mathbf{E}[(X-\mu_X)(Y-\mu_Y)] \quad \textbf{(væntigildisregla)} \\
&= a\mathrm{Cov}[X,Y]
\end{aligned}$$

Skilgreining: ***Samdreifni*** **slembistærða X og Y er**

$$\mathrm{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

Reglur: **Fyrir $a \in \mathbb{R}$ og $n, m \in \mathbb{N}$ gildir að**

(1) $\mathrm{Cov}[aX, Y] = a\,\mathrm{Cov}[X,Y]$

(2) $\mathrm{Cov}[X+Y, Z] = \mathrm{Cov}[X,Z] + \mathrm{Cov}[Y,Z]$

Sönnun á (2): **Notum $\mu_{X+Y} = \mu_X + \mu_Y$ í fyrsta skrefi í eftirfarandi og væntigildisreglu í fjórða,**

$$\begin{aligned}
\mathrm{Cov}[X+Y,Z] &= \mathbf{E}\big[\big((X+Y)-(\mu_X+\mu_Y)\big)\big(Z-\mu_Z\big)\big] \\
&= \mathbf{E}\big[\big((X-\mu_X)+(Y-\mu_Y)\big)\big(Z-\mu_Z\big)\big] \\
&= \mathbf{E}[(X-\mu_X)(Z-\mu_Z)+(Y-\mu_Y)(Z-\mu_Z)] \\
&= \mathbf{E}[(X-\mu_X)(Z-\mu_Z)] + \mathbf{E}[(Y-\mu_Y)(Z-\mu_Z)] \\
&= \mathrm{Cov}[X,Z] \qquad\qquad\qquad + \mathrm{Cov}[Y,Z]
\end{aligned}$$

Skilgreining: ***Samdreifni*** **slembistærða X og Y er**

$$\mathrm{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

Reglur: **Fyrir $a \in \mathbb{R}$ og $n, m \in \mathbb{N}$ gildir að**

(1) $\quad \mathrm{Cov}[aX, Y] = a\,\mathrm{Cov}[X,Y]$

(2) $\quad \mathrm{Cov}[X+Y, Z] = \mathrm{Cov}[X,Z] + \mathrm{Cov}[Y,Z]$

(3) $\quad \mathrm{Cov}\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}[X_i, Y_j]$

Sönnun á (3): **Setjum $Z = \sum_{j=1}^{m} Y_j$ og notum (2),**

$$\begin{aligned}
\mathrm{Cov}\left[\sum_{i=1}^{n} X_i, Z\right] &= \mathrm{Cov}\left[\sum_{i=1}^{n-1} X_i, Z\right] + \mathrm{Cov}[X_n, Z] \\
&= \ldots = \sum_{i=1}^{n} \mathrm{Cov}[X_i, Z] \\
&= \sum_{i=1}^{n} \left(\mathrm{Cov}\left[X_i, \sum_{j=1}^{m-1} Y_j\right] + \mathrm{Cov}[X_i, Y_m]\right) \\
&= \ldots = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}[X_i, Y_j].
\end{aligned}$$

Skilgreining: ***Samdreifni*** **slembistærða X og Y er**

$$\mathrm{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

Reglur: **Fyrir $a \in \mathbb{R}$ og $n, m \in \mathbb{N}$ gildir að**

(1) $\mathrm{Cov}[aX, Y] = a\,\mathrm{Cov}[X,Y]$

(2) $\mathrm{Cov}[X+Y, Z] = \mathrm{Cov}[X,Z] + \mathrm{Cov}[Y,Z]$

(3) $$\mathrm{Cov}\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}[X_i, Y_j]$$

(4) $$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + 2\sum_{i<j} \mathrm{Cov}[X_i, X_j]$$

Sönnun á (4): **Samkvæmt (3) gildir að**

$$\mathrm{Cov}\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}[X_i, X_j]$$

Skiptum hægri hliðinni miðað við $i = j$ og $i \neq j$,

$$\mathrm{Cov}\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right] = \sum_{i=1}^{n} \mathrm{Cov}[X_i, X_i] + \sum_{i\neq j} \mathrm{Cov}[X_i, X_j]$$

Notum $\mathrm{Cov}[X,X] = \mathrm{Var}[X]$ til að fá

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + \sum_{i\neq j} \mathrm{Cov}[X_i, X_j]$$

Notum $\mathrm{Cov}[X_i, X_j] = \mathrm{Cov}[X_j, X_i]$ til að fá

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + 2\sum_{i<j} \mathrm{Cov}[X_i, X_j]$$

Skilgreining: ***Samdreifni*** **slembistærða** X **og** Y **er**

$$\mathrm{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

Reglur: **Fyrir** $a \in \mathbb{R}$ **og** $n, m \in \mathbb{N}$ **gildir að**

(1) $\mathrm{Cov}[aX, Y] = a\,\mathrm{Cov}[X,Y]$

(2) $\mathrm{Cov}[X+Y, Z] = \mathrm{Cov}[X,Z] + \mathrm{Cov}[Y,Z]$

(3) $$\mathrm{Cov}\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}[X_i, Y_j]$$

(4) $$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + 2\sum_{i<j} \mathrm{Cov}[X_i, X_j]$$

Regla: **Ef** X **og** Y **eru óháðar gildir** $\mathrm{Cov}[X,Y] = 0$.

Sönnun: **Notum reiknireglunna**

$$\mathrm{Cov}[X,Y] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$$

og regluna fyrir væntigildi sem segir að ef X **og** Y **eru óháðar gildir** $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$**:**

$$\mathrm{Cov}[X,Y] = \mathbf{E}[X]\mathbf{E}[Y] - \mathbf{E}[X]\mathbf{E}[Y] = 0$$

Skilgreining: ***Samdreifni*** **slembistærða** X **og** Y **er**

$$\mathrm{Cov}[X,Y] := \mathbf{E}[(X-\mu_X)(Y-\mu_Y)]$$

Reglur: **Fyrir** $a \in \mathbb{R}$ **og** $n, m \in \mathbb{N}$ **gildir að**

(1) $\mathrm{Cov}[aX, Y] = a\,\mathrm{Cov}[X,Y]$

(2) $\mathrm{Cov}[X+Y, Z] = \mathrm{Cov}[X,Z] + \mathrm{Cov}[Y,Z]$

(3) $$\mathrm{Cov}\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}[X_i, Y_j]$$

(4) $$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i] + 2\sum_{i<j} \mathrm{Cov}[X_i, X_j]$$

Regla: **Ef** X **og** Y **eru óháðar gildir** $\mathrm{Cov}[X,Y] = 0$.

Þessi regla – ásamt reglu (4) – hefur nú í för með sér eftirfarandi reglu sem sett var fram á Aðalsíðu 17 (og hliðarsíðu 23):

Regla: **Ef** $X_1, \ldots, X_n$ **eru óháðar gildir að**

$$\mathrm{Var}[X_1 + \ldots + X_n] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n].$$

Hér er Höfuðmarkgildisreglunni beitt – A33 – L31

Höfuðmarkgildisreglan: Fyrir óháðar stærðir með sömu dreifingu gildir

$$X_1 + \ldots + X_n \approx \mathrm{N}(n\mu, n\sigma^2) \qquad \text{þegar } n \text{ stórt.}$$

Dæmi: Byggja á 1000 íbúða hverfi. Vitað er að bílafjöldi á dæmigerða fjölskyldu í svona hverfi hefur eftirfarandi dreifingu:

20% á engan bíl, 70% á einn bíl, 10% á tvo bíla.

Spurning: Hvað þarf stóra bílageymslu til að hún dugi fyrir hverfið með u.þ.b. 95% líkum?

Lausn: Látum X_i tákna bílafjölda á íbúð númer i. Gefum okkur að $X_1, \ldots, X_{1000}$ séu óháðar stærðir og

$$\mathbf{P}(X_i = 0) = 0,2 \qquad \mathbf{P}(X_i = 1) = 0,7 \qquad \mathbf{P}(X_i = 2) = 0,1$$

Fjöldi bílastæða k á að uppfylla

$$\mathbf{P}(X_1 + \ldots + X_{1000} \leqslant k) \approx 0,95$$

Samkvæmt Höfuðmarkgildisreglunni er

$$\mathbf{P}(X_1 + \ldots + X_{1000} \leqslant k) \approx \Phi\left(\frac{k - \mu\, 1000}{\sigma\sqrt{1000}}\right)$$

Nú er $\mu = \mathbf{E}[X_i] = 0 \cdot 0,2 + 1 \cdot 0,7 + 2 \cdot 0.1 = 0,9$ og

$$\sigma^2 = \mathrm{Var}[X_i] = \mathbf{E}[X_i^2] - \mu^2$$
$$= (0^2 \cdot 0,2 + 1^2 \cdot 0,7 + 2^2 \cdot 0,1) - 0,9^2 = 0,29$$

Samkvæmt töflu er $\Phi(1,645) \approx 0,95$ og því er

$$\frac{k - 900}{\sqrt{290}} \approx 1,645 \quad \text{sem gefur } k \approx 900 + 17 \cdot 1,645 \approx 928$$

Á þessari opnu er reglum af A33-36 beitt – L35

Dæmi úr gömlu dæmasafni: Verkfræðinemar setjast við glugga í VR 2 og fylgjast með umferð um Suðurgötu í báðar áttir. Til norðurs aka að jafnaði 8 bílar á mínútu og 4 bílar á mínútu til suðurs. Umferðartíðni er stöðug og bílarnir hafa ekki áhrif hver á annan [lýsing á óháðum Poissonferlum, HÞ].

a) Finnið væntanlegt gildi og fervik [dreifni, HÞ] tímans sem líður þar til 4. bíll ekur fram hjá ljósastaurnum sem miðað er við.

b) Finnið líkurnar á að bílarnir 4 í a) lið séu allir á suðurleið.

c) Finnið líkurnar á að færri en 220 bílar aki til norðurs næstu 30 mínúturnar eftir að talning hefst.

Lausn: Eftirfarandi ferli eru óháð Poisson-ferli:

$L_t =$ fjöldi bíla á leið norður á bilinu $[0,t]$ $\quad (a=8)$

$M_t =$ fjöldi bíla a leið suður á bilinu $[0,t]$ $\quad (b=4)$

a) $\mathbf{E}[X_1+X_2+X_3+X_4] = 4\cdot\frac{1}{c} = 4\cdot\frac{1}{8+4} = \frac{1}{3}$

$\mathrm{Var}[X_1+X_2+X_3+X_4] = 4\cdot\frac{1}{c^2} = 4\cdot\frac{1}{(8+4)^2} = \frac{4}{144} = \frac{1}{36}$

b) $\mathbf{P}(I_1=I_2=I_3=I_4=1) = p\cdot p\cdot p\cdot p = \left(\frac{4}{8+4}\right)^4 = \frac{1}{81}$

c) Vegna þess að $L_{30}\sim \mathrm{Poi}(30\cdot 8)$ og $30\cdot 8 = 240 \geqslant 10$ höfum við að $L_{30}\approx \mathrm{N}(240,240)$ sem gefur nálgunina

$$\mathbf{P}(L_{30}\leqslant 219) \approx \Phi\left(\frac{219-240}{\sqrt{240}}\right)$$

$$= \Phi\big(-1,35\big) = 1-\Phi\big(1,35\big) = 1-0,91 = 0,09 = 9\%$$

Framhald á VR 2 dæminu:

---

c) Finnið líkurnar á að færri en 220 bílar aki til norðurs næstu 30 mínúturnar eftir að talning hefst.

Við leystum þetta svona á síðunni hér á undan:

Eftirfarandi ferli er Poisson-ferli með tíðni $a = 8$:

$$L_t = \text{fjöldi bíla á leið norður á bilinu } [0, t].$$

Vegna þess að $L_{30} \sim \mathrm{Poi}(30 \cdot 8)$ og $30 \cdot 8 = 240 \geqslant 10$ höfum við að $L_{30} \approx \mathrm{N}(240, 240)$ sem gefur nálgunina

$$\mathbf{P}(L_{30} \leqslant 219) \approx \Phi\big(\tfrac{219-240}{\sqrt{240}}\big)$$

$$= \Phi\big(-1,35\big) = 1 - \Phi\big(1,35\big) = 1 - 0,91 = 0,09 = 9\%$$

Hér er önnur nálgun: Látum S_n vera tímann þegar n^{ti} bíllin á leið norður fer fram hjá ljósastaurnum. Þá höfum við

$$\{L_{30} \leqslant 219\} = \{S_{220} > 30\}.$$

Nú er S_n summa n óháðra $\mathrm{Exp}(8)$ stærða.
Leyfum okkur að ætla að $n = 220$ sé nægilega mikill fjöldi til að nota megi Höfuðmarkgidissetninguna:
Væntigildi $\mathrm{Exp}(8)$ stærðar er $1/8$ og dreifnin er $1/8^2$.
Við fáum

$$\mathbf{P}(L_{30} \leqslant 219) = \mathbf{P}(S_{220} > 30) = 1 - \mathbf{P}(S_{220} \leqslant 30)$$

$$\approx 1 - \Phi\big(\tfrac{30 - 220 \cdot 1/8}{\sqrt{220} \cdot 1/8}\big) = 1 - \Phi\big(1,35\big) = \ldots = 9\%$$