
Ilkka Mellin
Todennäköisyyslaskenta

**Osa 2: Satunnaismuuttujat ja
todennäköisyysjakaumat**
**Moniulotteiset satunnaismuuttujat ja
jakaumat**

Moniulotteiset satunnaismuuttajat ja todennäköisyysjakaumat

>> Kaksiulotteiset todennäköisyysjakaumat

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Kovarianssi ja korrelaatio

Ehdolliset jakaumat ja odotusarvot

Satunnaistekijöiden väliset riippuvuudet

- *Yhden* satunnaismuuttujan todennäköisyysjakaumat kuvaavat useimpia satunnaisilmiötä vain rajoitetusti.
- Satunnaisilmiöihin liittyy tavallisesti **useita satunnaisia tekijöitä**, joiden väliset **riippuvuudet** ovat varsinaisen mielenkiinnon kohteina.
- Useiden satunnaisten tekijöiden välisten riippuvuuksien *mallintaminen* vaatii tekijöihin liittyvien satunnaismuuttujien **yhteisjakauman** tarkastelua.

Esimerkkejä riippuvuustarkasteluista

- Miten työttömyysaste Suomessa (% työvoimasta) *riippuu* BKT:n kasvuvauhdista, viennin volyyymistä ja BKT:n kasvuvauhdista muissa EU-maissa ja USA:ssa?
- Miten alkoholin kokonaiskulutus (1 *per capita* vuodessa) *riippuu* alkoholijuomien hintatasosta, käytettävissä olevista tuloista ja alkoholin saatavuudesta?
- Miten todennäköisyys sairastua keuhkosityöpään (p) *riippuu* tupakoinnin määrästä ja kestosta?
- Miten vehnän sato (t/ha) *riippuu* kesän keskilämpötilasta, sademäärästä, maan muokkaustavoista, lannoituksesta ja tuholaisten torjunnasta?

Kaksiulotteiset satunnaismuuttujat

- Olkoot X ja Y satunnaismuuttujia, joiden otosavaruudet ovat R ja S .

- Tällöin

$$X : S \rightarrow$$

$$Y : R \rightarrow$$

- Olkoon $R \times S$ otosavaruuksien R ja S *karteesinen tulo*:

$$R \times S = \{(r, s) \mid r \in R, s \in S\}$$

- Satunnaismuuttujien X ja Y *järjestetty pari* (X, Y) määrittelee **kaksiulotteisen satunnaismuuttujan**:

$$(X, Y) : S \times R \rightarrow \mathbb{R}^2$$

Diskreetit kaksiulotteiset satunnaismuuttujat ja niiden jakaumat

- Olkoot X ja Y *diskreettejä* satunnaismuuttujia.
- Tällöin järjestetty pari (X, Y) määrittelee **diskreetin kaksiulotteisen satunnaismuuttujan**.
- Diskreetti kaksiulotteinen satunnaismuuttuja (X, Y) määrittelee **diskreetin kaksiulotteisen todennäköisyysjakauman**, jota kutsutaan satunnaismuuttujien X ja Y yhteisjakaumaksi.

Diskreetit kaksiulotteiset jakaumat: Pistetodennäköisyysfunktio

- Reaaliarvoinen funktio

$$f_{XY} : \mathbb{R}^2 \rightarrow \mathbb{R}$$

määrittelee *diskreettien satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktion*, jos seuraavat ehdot pätevät:

$$(1) \quad f_{XY}(x, y) \geq 0 \text{ kaikille } x \text{ ja } y$$

$$(2) \quad \sum_x \sum_y f_{XY}(x, y) = 1$$

$$(3) \quad \Pr(X = x \text{ ja } Y = y) = f_{XY}(x, y)$$

Diskreetit kaksiulotteiset jakaumat: Tapahtumien todennäköisyydet

- Olkoon

$$f_{XY} : \mathbb{R}^2 \rightarrow \mathbb{R}$$

diskreettien satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio.

- Olkoon

$$A \subset \mathbb{R}^2$$

jokin tapahtuma.

- Tällöin

$$\Pr((X, Y) \in A) = \sum_{(x, y) \in A} f_{XY}(x, y)$$

Diskreetit kaksiulotteiset jakaumat: Symmetriset todennäköisyyskentät

- Olkoon (x_i, y_i) , $i = 1, 2, \dots, n$ erillisten tason pisteiden muodostama *diskreetti* pistejoukko.
- Määritellään *diskreetin kaksiulotteisen jakauman pistetodennäköisyysfunktio* kaavalla

$$f_{XY}(x_i, y_i) = \Pr(X = x_i \text{ ja } Y = y_i) = \frac{1}{n}, i = 1, 2, \dots, n$$

- Tällöin satunnaismuuttuja (X, Y) määrittelee **symmetrisen todennäköisyyskentän**, jossa kaikki alkeistapahtumat

$$(x_i, y_i), i = 1, 2, \dots, n$$

ovat *yhtä todennäköisiä*.

Diskreetit kaksiulotteiset jakaumat:

1. esimerkki nopanheitosta 1/6

- Heitetään *virheetöntä* noppaa kaksi kertaa.
- Määritellään *satunnaismuuttujat* X ja Y :
 $X = \text{tulos (silmäluku) 1. heitosta}$
 $Y = \text{tulos (silmäluku) 2. heitosta}$
- Voimme olettaa, että 2. heiton tulos on *riippumaton* 1. heiton tuloksesta (ja kääntäen).
- Satunnaismuuttujien X ja Y mahdolliset arvot:
 $X: \{1, 2, 3, 4, 5, 6\}$
 $Y: \{1, 2, 3, 4, 5, 6\}$
- Muodostetaan satunnaismuuttujien X ja Y *yhteisjakauma*.

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

1. esimerkki nopanheitosta 2/6

- Kahden nopanheiton tulosvaihtoehdot voidaan esittää seuraavana taulukkona:

Tulos 1. nopanheitosta	Mahdolliset tulokset 2. nopanheitosta					
1	1	2	3	4	5	6
2	1	2	3	4	5	6
3	1	2	3	4	5	6
4	1	2	3	4	5	6
5	1	2	3	4	5	6
6	1	2	3	4	5	6

- Siten satunnaismuuttujien X ja Y järjestetty pari (X, Y) määrittelee *diskreetin kaksiulotteisen satunnaismuuttujan*, jonka arvoina on 36 lukuparia

$$(x, y) ; x = 1, 2, 3, 4, 5, 6 ; y = 1, 2, 3, 4, 5, 6$$

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

1. esimerkki nopanheitosta 3/6

- Koska noppa oletettiin *virheettömäksi* ja heittojen tulokset oletettiin *riippumattomiksi*, on luontevaa ajatella, että kahden nopanheiton tulosten muodostamat 36 tulosvaihtoehtoa

$$(x, y) ; x = 1, 2, 3, 4, 5, 6 ; y = 1, 2, 3, 4, 5, 6$$

ovat *yhtä todennäköisiä*.

- Tällöin satunnaismuuttujien X ja Y yhteisjakauman *pistetodennäköisyysfunktio* saa positiiviset arvot

$$f_{XY}(x, y) = \Pr(X = x, Y = y) = \Pr(X = x) \Pr(Y = y) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

kun

$$x = 1, 2, 3, 4, 5, 6$$

$$y = 1, 2, 3, 4, 5, 6$$

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

1. esimerkki nopanheitosta 4/6

- Satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio

$$f_{XY}(x, y) = \Pr(X = x \text{ ja } Y = y) = \frac{1}{36}$$

$$x = 1, 2, 3, 4, 5, 6; y = 1, 2, 3, 4, 5, 6$$

voidaan esittää seuraavana taulukkona:

2. heiton silmäluku y	6	1/36	1/36	1/36	1/36	1/36	1/36
	5	1/36	1/36	1/36	1/36	1/36	1/36
	4	1/36	1/36	1/36	1/36	1/36	1/36
	3	1/36	1/36	1/36	1/36	1/36	1/36
	2	1/36	1/36	1/36	1/36	1/36	1/36
	1	1/36	1/36	1/36	1/36	1/36	1/36
		1	2	3	4	5	6

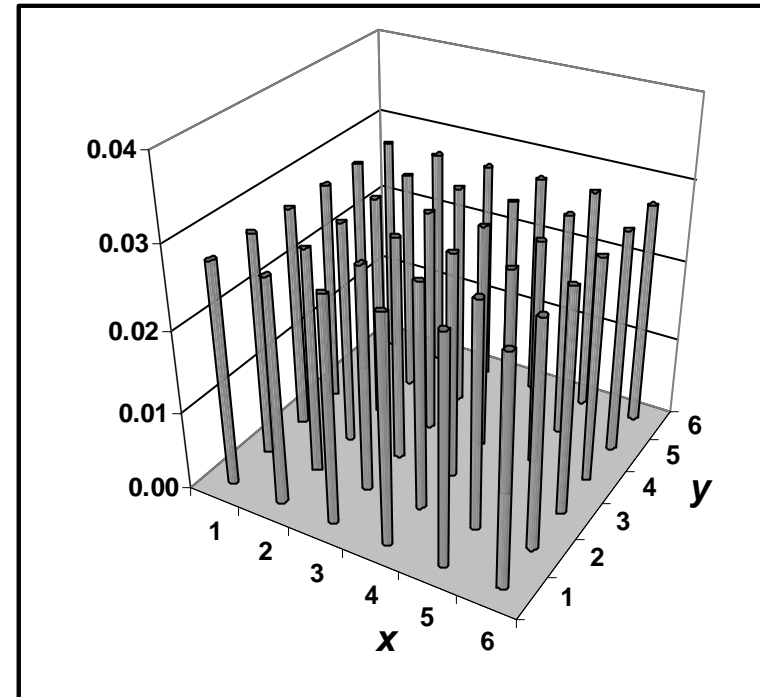
1. heiton silmäluku x

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

1. esimerkki nopanheitosta 5/6

- Heitetään *virheetöntä* noppaa kaksi kertaa:
 X = tulos 1. heitosta
 Y = tulos 2. heitosta
- Kuva oikealla esittää satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktiota.
- Kuvassa on 36 pylvästä, joista jokaisen korkeus $1/36$.

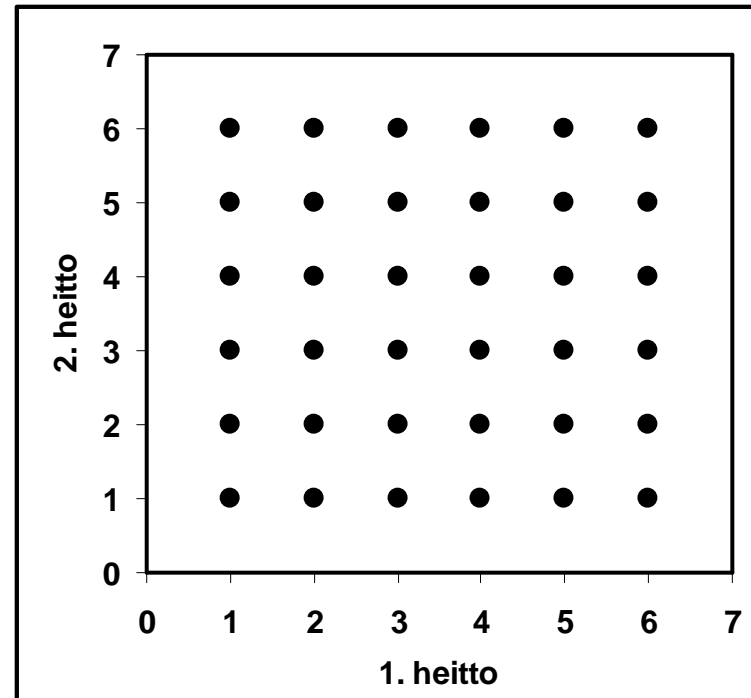


Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

1. esimerkki nopanheitosta 6/6

- Heitetään *virheetöntä* noppaa kaksi kertaa:
 X = tulos 1. heitosta
 Y = tulos 2. heitosta
- Kuva oikealla havainnollistaa satunnaismuuttujien X ja Y yhteisjakaumaa.
- Kuvassa on 36 pistettä, joista jokaisen todennäköisyys on $1/36$.



Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 1/9

- Heitetään *virheetöntä* noppaa kaksi kertaa.
- Määritellään *satunnaismuuttujat* X , Y ja Z :
 - $X =$ tulos (silmäluku) 1. heitosta
 - $Y =$ tulos (silmäluku) 2. heitosta
 - $Z = X + Y =$ silmälukujen summa
- Voimme olettaa, että 2. heiton tulos on *riippumaton* 1. heiton tuloksesta (ja kääntäen).
- Satunnaismuuttujien X , Y ja Z mahdolliset arvot:
 - $X: \{1, 2, 3, 4, 5, 6\}$
 - $Y: \{1, 2, 3, 4, 5, 6\}$
 - $Z: \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
- Määrätään satunnaismuuttujien X ja Z yhteisjakauma.

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 2/9

- Muodostetaan ensin *summamuuttujan* $Z = X + Y$ jakauma.
- Muodostetaan sitä varten *aputaulukko*, joka esittää kaikkia mahdollisia tapoja, joilla nopanheittojen silmälukujen summa $Z = X + Y$ voi syntyä:

Silmälukujen summat $z = x + y$

2. heiton silmäluku y	6	7	8	9	10	11	12
	5	6	7	8	9	10	11
	4	5	6	7	8	9	10
	3	4	5	6	7	8	9
	2	3	4	5	6	7	8
	1	2	3	4	5	6	7
		1	2	3	4	5	6

1. heiton silmäluku x

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 3/9

- Aputaulukosta voidaan suoraan lukea 1. ja 2. nopanheiton silmälukujen *summan*

$$Z = X + Y$$

todennäköisyysjakauma:

Silmälukujen summat $z = x + y$ ja niiden todennäköisyydet

2	3	4	5	6	7	8	9	10	11	12
1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

- Esimerkki:

Summa 5 voi syntyä kahden nopanheiton tuloksena 4:llä eri tavalla:

$$5 = 1 + 4 = 2 + 3 = 3 + 2 = 4 + 1$$

joten todennäköisyys saada summaksi 5 on 4/36.

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 4/9

- 1. nopanheiton tulos ja 1. ja 2. nopanheiton tulosten kaikki mahdolliset summat voidaan esittää seuraavana taulukkona:

Tulos 1. nopanheitosta	Mahdolliset summat 1. ja 2. nopanheiton tuloksista					
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

- Siten satunnaismuuttujien X ja $Z = X + Y$ järjestetty pari (X, Z) määrittelee *diskreetin kaksiulotteisen satunnaismuuttujan*, jonka arvoina on 66 lukuparia

$$(x, z) ; x = 1, 2, 3, 4, 5, 6 ; z = 2, 3, 4, 5, 6, 7, 8, 10, 11, 12$$

Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 5/9

- Koska noppa oletettiin *virheettömäksi* ja heittojen tulokset oletettiin *riippumattomiksi*, on luontevaa ajatella, että 1. heiton tulos ja 1. ja 2. heiton tulosten *mahdollisten* summien muodostamat 36 tulosvaihtoehtoa

$$(x, z) ; x = 1, 2, 3, 4, 5, 6 ; z = x + y ; y = 1, 2, 3, 4, 5, 6$$

ovat *yhtä todennäköisiä*.

- Tällöin satunnaismuuttujien X ja Y yhteisjakauman *pistetodennäköisyysfunktion* saa positiiviset arvot

$$f_{XZ}(x, z) = \Pr(X = x \text{ ja } Z = z) = \frac{1}{36}$$

kun

$$x = 1, 2, 3, 4, 5, 6$$

$$z = x + y$$

$$y = 1, 2, 3, 4, 5, 6$$

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 6/9

- Satunnaismuuttujien X ja $Z = X + Y$ yhteisjakauman pistetodennäköisyysfunktio voidaan esittää seuraavana taulukkona:

Silmälukujen summa z	12	0	0	0	0	0	1/36
	11	0	0	0	0	1/36	1/36
	10	0	0	0	1/36	1/36	1/36
	9	0	0	1/36	1/36	1/36	1/36
	8	0	1/36	1/36	1/36	1/36	1/36
	7	1/36	1/36	1/36	1/36	1/36	1/36
	6	1/36	1/36	1/36	1/36	1/36	0
	5	1/36	1/36	1/36	1/36	0	0
	4	1/36	1/36	1/36	0	0	0
	3	1/36	1/36	0	0	0	0
	2	1/36	0	0	0	0	0
		1	2	3	4	5	6
1. nopan silmäluku x							

Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 7/9

- Esimerkkejä:

- (i) Oletetaan, että 1. nopalla on saatu 2. Tällöin silmälukujen summaksi *ei voi tulla* 10, joten

$$\Pr(X = 1 \text{ ja } Z = 10) = 0$$

- (ii) Oletetaan, että 1. nopalla on saatu 6. Tällöin silmälukujen summaksi *ei voi tulla* 3, joten

$$\Pr(X = 6 \text{ ja } Z = 3) = 0$$

- (iii) Oletetaan, että 1. nopalla on saatu 2. Tällöin silmälukujen summaksi *voi tulla* 3, 4, 5, 6, 7 tai 8, joten

$$\Pr(X = 1 \text{ ja } Z = z) = 1/36 ; z = 3, 4, 5, 6, 7, 8$$

- (iv) Oletetaan, että 1. nopalla on saatu 6. Tällöin silmälukujen summaksi *voi tulla* 7, 8, 9, 10, 11 tai 12, joten

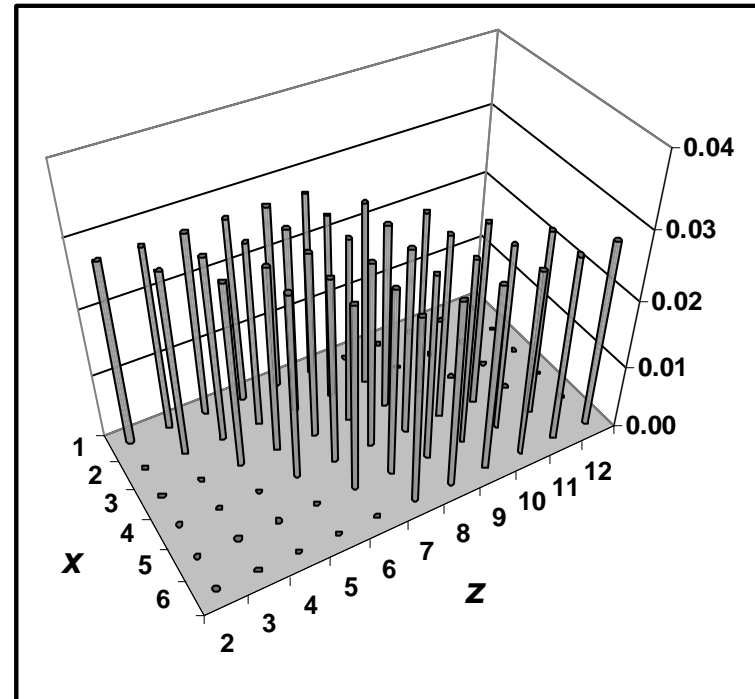
$$\Pr(X = 6 \text{ ja } Z = z) = 1/36 ; z = 7, 8, 9, 10, 11, 12$$

Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 8/9

- Heitetään *virheetöntä* noppaa kaksi kertaa:
 X = tulos 1. heitosta
 Y = tulos 2. heitosta
 $Z = X + Y$
- Kuva oikealla esittää satunnaismuuttujien X ja Z yhteisjakauman pistetodennäköisyysfunktiota.
- Kuvassa on 36 pylvästä, joista jokaisen korkeus on $1/36$.

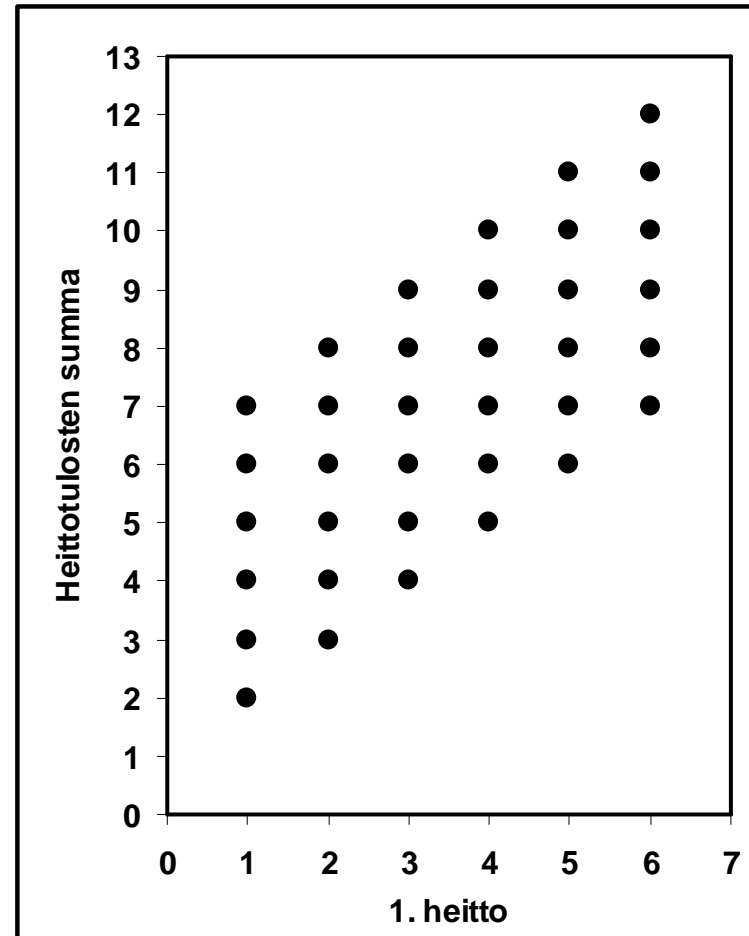


Kaksiulotteiset todennäköisyysjakaumat

Diskreetit kaksiulotteiset jakaumat:

2. esimerkki nopanheitosta 9/9

- Heitetään *virheetöntä* noppaa kaksi kertaa:
 $X = \text{tulos 1. heitosta}$
 $Y = \text{tulos 2. heitosta}$
 $Z = X + Y$
- Kuva oikealla havainnollistaa satunnaismuuttujien X ja Z yhteisjakaumaa.
- Kuvassa on 36 pistettä, joista jokaisen todennäköisyys on $1/36$.



Jatkuvat kaksiulotteiset satunnaismuuttajat ja niiden jakaumat

- Olkoot X ja Y *jatkuvia* satunnaismuuttujia.
- Tällöin järjestetty pari (X, Y) määrittelee **jatkuvan kaksiulotteisen satunnaismuuttujan**.
- Kaksiulotteinen jatkuva satunnaismuuttuja (X, Y) määrittelee **jatkuvan kaksiulotteisen todennäköisyysjakauman**, jota kutsutaan satunnaismuuttujien X ja Y yhteisjakaumaksi.

Jatkuvat kaksiulotteiset jakaumat: Tiheysfunktio

- Reaaliarvoinen jatkuva funktio

$$f_{XY} : \mathbb{R}^2 \rightarrow \mathbb{R}$$

määrittelee *jatkuvien satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktion*, jos seuraavat ehdot pätevät:

$$(1) \quad f_{XY}(x, y) \geq 0 \text{ kaikille } x \text{ ja } y$$

$$(2) \quad \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{XY}(x, y) dy dx = 1$$

$$(3) \quad \Pr(a \leq X \leq b \text{ ja } c \leq Y \leq d) = \int_a^b \int_c^d f_{XY}(x, y) dy dx$$

Kaksiulotteiset todennäköisyysjakaumat

Kaksinkertaiset integraalit: Integrointijärjestys

- Huomautus:

Käytämme *kaksinkertaisten integraalien* yhteydessä seuraavaa sopimusta integrointijärjestyksestä:

$$\int_a^b \int_c^d f_{XY}(x, y) dy dx = \int_a^b \left(\int_c^d f_{XY}(x, y) dy \right) dx$$

Jatkuvat kaksiulotteiset jakaumat: Tapahtumien todennäköisyydet

- Olkoon

$$f_{XY} : \mathbb{R}^2 \rightarrow \mathbb{R}$$

jatkuvien satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio.

- Olkoon

$$A \subset \mathbb{R}^2$$

jokin tapahtuma.

- Tällöin

$$\Pr((X, Y) \in A) = \iint_A f_{XY}(x, y) dy dx$$

Kaksiulotteiset todennäköisyysjakaumat

Kaksiulotteisten jakaumien kertymäfunktiot

- Olkoon (X, Y) satunnaismuuttujien X ja Y muodostama *järjestetty pari*.
- Satunnaismuuttujien X ja Y yhteisjakauman **kertymäfunktio** F_{XY} määritellään kaavalla:

$$F_{XY}(x, y) = \Pr(X \leq x \text{ ja } Y \leq y)$$

Diskreettien kaksiulotteisten jakaumien kertymäfunktiot

- Olkoon (X, Y) *diskreetti* kaksiulotteinen satunnaismuuttuja.
- Olkoon $\{x_1, x_2, x_3, \dots\}$ satunnaismuuttujan X tulosvaihtoehtojen eli arvojen joukko.
- Olkoon $\{y_1, y_2, y_3, \dots\}$ satunnaismuuttujan Y tulosvaihtoehtojen eli arvojen joukko.
- **Diskreetin kaksiulotteisen jakauman kertymäfunktio on**

$$\begin{aligned} F_{XY}(x, y) &= \Pr(X \leq x \text{ ja } Y \leq y) \\ &= \sum_{x_i \leq x} \sum_{y_i \leq y} f_{XY}(x_i, y_i) \end{aligned}$$

jossa f_{XY} satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio.

Jatkuvien kaksiulotteisten jakaumien kertymäfunktiot

- Olkoon (X, Y) *jatkuva* kaksiulotteinen satunnaismuuttuja.
- **Jatkuvan kaksiulotteisen jakauman kertymäfunktio** on

$$F_{XY}(x, y) = \Pr(X \leq x \text{ ja } Y \leq y)$$

$$= \int_{-\infty}^x \int_{-\infty}^y f_{XY}(u, v) dv du$$

jossa f_{XY} satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio.

Jatkuvan kaksiulotteisen jakauman tiheysfunktion ja kertymäfunktion yhteys

- Olkoon (X, Y) *jatkuva* kaksiulotteinen satunnaismuuttuja.
- Olkoon $F_{XY}(x, y)$ satunnaismuuttujien X ja Y yhteisjakauman *kertymäfunktio*.

- Jos *derivaatta*

$$\frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} = f_{XY}(x, y)$$

on olemassa ja on jatkuva, funktio $f_{XY}(x, y)$ on satunnaismuuttujien X ja Y yhteisjakauman *tiheysfunktio*.

Moniulotteiset satunnaismuuttujat ja todennäköisyysjakaumat

Kaksiulotteiset todennäköisyysjakaumat

>> Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Kovarianssi ja korrelaatio

Ehdolliset jakaumat ja odotusarvot

Diskreetin kaksiulotteisen jakauman reunajakaumat

- Olkoon $f_{XY}(x, y)$ *diskreetin* kaksiulotteisen jakauman pistetodennäköisyysfunktio.

- Satunnaismuuttujan X **reunajakauman pistetodennäköisyysfunktio** on

$$f_X(x) = \Pr(X = x) = \sum_y f_{XY}(x, y)$$

- Satunnaismuuttujan Y **reunajakauman pistetodennäköisyysfunktio** on

$$f_Y(y) = \Pr(Y = y) = \sum_x f_{XY}(x, y)$$

- Satunnaismuuttujien X ja Y reunajakaumat *yhtyvät* satunnaismuuttujien X ja Y todennäköisyysjakaumiin.

Diskreetin 2-ulotteisen jakauman reunajakaumat:

1. esimerkki nopanheitosta 1/5

- Heitetään *virheetöntä* noppaa kaksi kertaa.
- Määritellään *satunnaismuuttujat* X ja Y :
 $X = \text{tulos (silmäluku) 1. heitosta}$
 $Y = \text{tulos (silmäluku) 2. heitosta}$
- Satunnaismuuttujien X ja Y mahdolliset arvot:
 $X: \{1, 2, 3, 4, 5, 6\}$
 $Y: \{1, 2, 3, 4, 5, 6\}$
- Määrätään satunnaismuuttujien X ja Y *reunajakaumat*.

Diskreetin 2-ulotteisen jakauman reunajakaumat:

1. esimerkki nopanheitosta 2/5

- Satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio:

$$\Pr(X = x \text{ ja } Y = y) = f_{XY}(x, y) ; x = 1, 2, \dots, 6 ; y = 1, 2, \dots, 6$$

1. heiton silmäluku x

	1	2	3	4	5	6
6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36

2. heiton silmäluku y

Diskreetin 2-ulotteisen jakauman reunajakaumat:

1. esimerkki nopanheitosta 3/5

- Satunnaismuuttujien X ja Y reunajakaumien pistetodennäköisyysfunktioit saadaan määräämällä yhteisjakauman todennäköisyydet antavassa taulukossa rivi- ja sarakesummat.
- Satunnaismuuttujan X reunajakauman pistetodennäköisyysfunktio on

$$f_X(x) = \Pr(X = x) = \sum_{y=1}^6 f_{XY}(x, y) = 6 \times \frac{1}{36} = \frac{1}{6}; x = 1, 2, 3, 4, 5, 6$$

koska

$$f_{XY}(x, y) = \frac{1}{36}; x = 1, 2, 3, 4, 5, 6; y = 1, 2, 3, 4, 5, 6$$

- Satunnaismuuttujan Y reunajakauman pistetodennäköisyysfunktio on

$$f_Y(y) = \Pr(Y = y) = \sum_{x=1}^6 f_{XY}(x, y) = 6 \times \frac{1}{36} = \frac{1}{6}; y = 1, 2, 3, 4, 5, 6$$

koska

$$f_{XY}(x, y) = \frac{1}{36}; x = 1, 2, 3, 4, 5, 6; y = 1, 2, 3, 4, 5, 6$$

Diskreetin 2-ulotteisen jakauman reunajakaumat:

1. esimerkki nopanheitosta 4/5

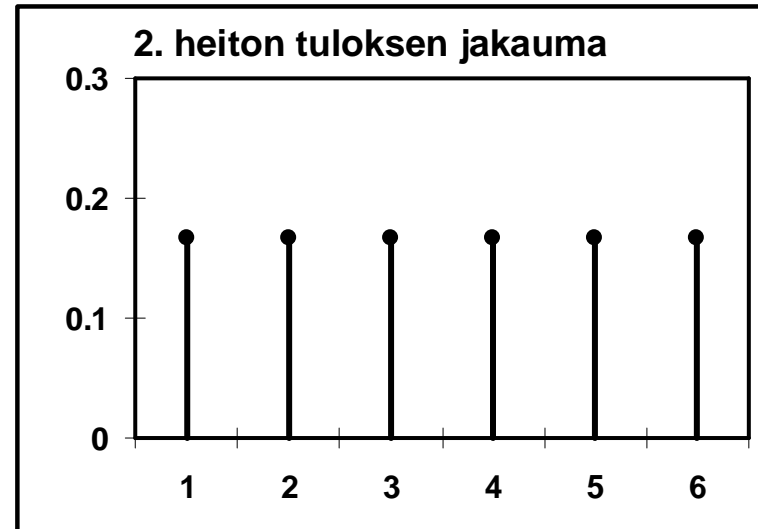
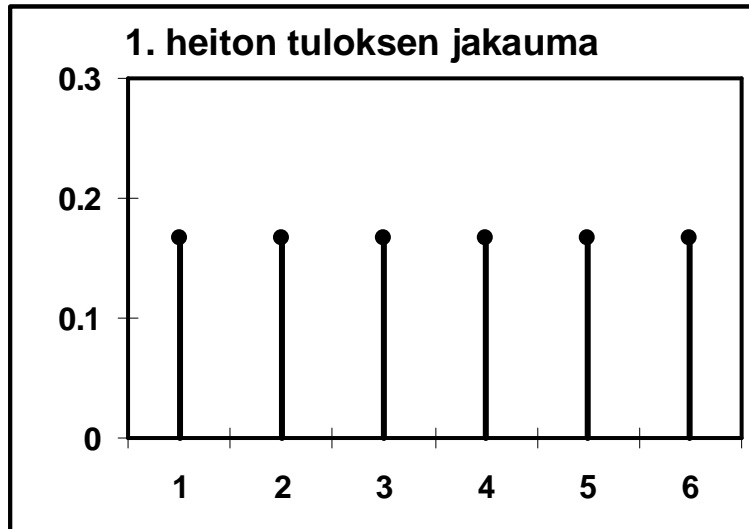
- Satunnaismuuttujien X ja Y yhteisjakauman ja reunajakaumien pistetodennäköisyysfunktiot:

1. heiton silmäluku x

	1	2	3	4	5	6	Yht
2. heiton silmäluku y							
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
Yht	1/6	1/6	1/6	1/6	1/6	1/6	1

Diskreetin 2-ulotteisen jakauman reunajakaumat:

1. esimerkki nopanheitosta 5/5



- Kuvat yllä esittävät satunnaismuuttujien X ja Y reunajakaumien pistetodennäköisyysfunktioita:

$X =$ tulos 1. heitosta

$Y =$ tulos 2. heitosta

Diskreetin 2-ulotteisen jakauman reunajakaumat: 2. esimerkki nopanheitosta 1/5

- Heitetään *virheetöntä* noppaa kaksi kertaa.
- Määritellään *satunnaismuuttujat* X , Y ja Z seuraavasti:
 - $X =$ tulos (silmäluku) 1. heitosta
 - $Y =$ tulos (silmäluku) 2. heitosta
 - $Z = X + Y =$ silmälukujen summa
- Satunnaismuuttujien X , Y ja Z mahdolliset arvot:
 - $X: \{1, 2, 3, 4, 5, 6\}$
 - $Y: \{1, 2, 3, 4, 5, 6\}$
 - $Z: \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
- Määrätään satunnaismuuttujien X ja Z *reunajakaumat*.

Diskreetin 2-ulotteisen jakauman reunajakaumat: 2. esimerkki nopanheitosta 2/5

- Satunnaismuuttujien X ja Z yhteisjakauman pistetodennäköisyysfunktio:

$$\Pr(X = x \text{ ja } Z = z) = f_{XZ}(x, z) ; x = 1, 2, \dots, 6 ; z = 2, 3, \dots, 12$$

1. nopan silmäluku x

	1	2	3	4	5	6
Silmälukujen summa z						
12	0	0	0	0	0	1/36
11	0	0	0	0	1/36	1/36
10	0	0	0	1/36	1/36	1/36
9	0	0	1/36	1/36	1/36	1/36
8	0	1/36	1/36	1/36	1/36	1/36
7	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	0
5	1/36	1/36	1/36	1/36	0	0
4	1/36	1/36	1/36	0	0	0
3	1/36	1/36	0	0	0	0
2	1/36	0	0	0	0	0

Diskreetin 2-ulotteisen jakauman reunajakaumat: 2. esimerkki nopanheitosta 3/5

- Satunnaismuuttujien X ja Y reunajakaumien pistetodennäköisyysfunktioit saadaan määräämällä yhteisjakauman todennäköisyydet antavassa taulukossa rivi- ja sarakesummat.
- Esimerkkejä:
 - (i) Satunnaismuuttujan X reunajakauman pistetodennäköisyys, kun $X = 4$:

$$\begin{aligned} f_X(4) &= \sum_{z=1}^{12} f_{XZ}(4, z) \\ &= 0 + 0 + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + 0 + 0 + 0 = \frac{1}{6} \end{aligned}$$

- (ii) Satunnaismuuttujan Z reunajakauman pistetodennäköisyys, kun $Z = 10$:

$$f_Z(10) = \sum_{x=1}^6 f_{XZ}(x, 10) = 0 + 0 + 0 + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{3}{36}$$

Diskreetin 2-ulotteisen jakauman reunajakaumat: 2. esimerkki nopanheitosta 4/5

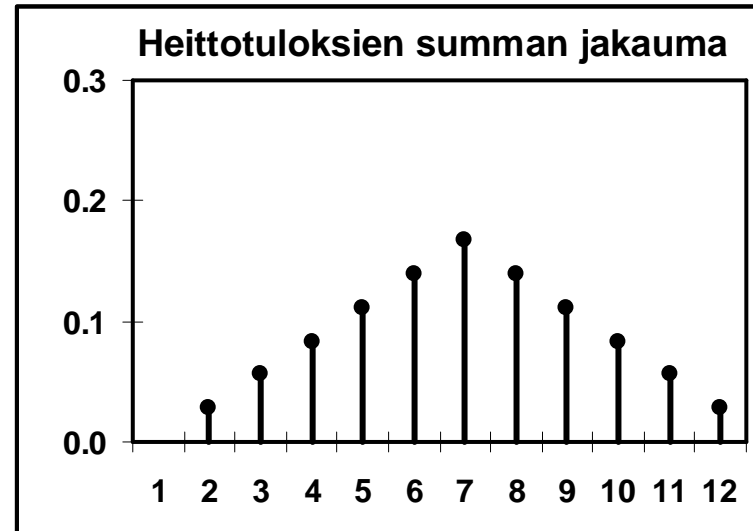
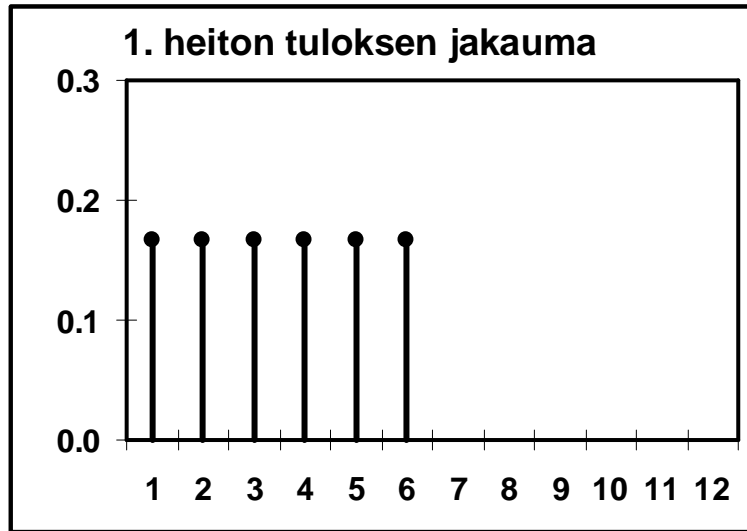
- Satunnaismuuttujien X ja $Z = X + Y$ yhteisjakauma ja reunajakaumat:

1. nopan silmäluku x

	1	2	3	4	5	6	Yht	
Silmälukujen summa z	12	0	0	0	0	0	1/36	1/36
	11	0	0	0	0	1/36	1/36	2/36
	10	0	0	0	1/36	1/36	1/36	3/36
	9	0	0	1/36	1/36	1/36	1/36	4/36
	8	0	1/36	1/36	1/36	1/36	1/36	5/36
	7	1/36	1/36	1/36	1/36	1/36	1/36	6/36
	6	1/36	1/36	1/36	1/36	1/36	0	5/36
	5	1/36	1/36	1/36	1/36	0	0	4/36
	4	1/36	1/36	1/36	0	0	0	3/36
	3	1/36	1/36	0	0	0	0	2/36
	2	1/36	0	0	0	0	0	1/36
	Yht	1/6	1/6	1/6	1/6	1/6	1/6	1

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Diskreetin 2-ulotteisen jakauman reunajakaumat: 2. esimerkki nopanheitosta 5/5



- Kuvat yllä esittävät satunnaismuuttujien X ja Z *reunajakaumien pistetodennäköisyysfunktioita*:

X = tulos 1. heitosta

Y = tulos 2. heitosta

$Z = X + Y$ = heittotulosten summa

Jatkuvan kaksiulotteisen jakauman reunajakaumat

- Olkoon $f_{XY}(x, y)$ *jatkuvan* kaksiulotteisen jakauman tiheysfunktio.
- Satunnaismuuttujan X **reunajakauman tiheysfunktio** on

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy$$

- Satunnaismuuttujan Y **reunajakauman tiheysfunktio** on

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx$$

- Satunnaismuuttujien X ja Y reunajakaumat *yhtyvät* satunnaismuuttujien X ja Y todennäköisyysjakaumiin.

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Satunnaismuuttujien riippumattomuus 1/2

- Oletukset:
 - (i) Olkoon satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyys- tai tiheysfunktio $f_{XY}(x, y)$.
 - (ii) Olkoon satunnaismuuttujan X reunajakauman pistetodennäköisyys- tai tiheysfunktio $f_X(x)$.
Olkoon satunnaismuuttujan Y reunajakauman pistetodennäköisyys- tai tiheysfunktio $f_Y(y)$.
- Määritelmä 1:
Satunnaismuuttujat X ja Y ovat **riippumattomia**, jos ja vain, jos

$$f_{XY}(x, y) = f_X(x)f_Y(y)$$

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Satunnaismuuttujien riippumattomuus 2/2

- Oletukset:
 - (i) Olkoon satunnaismuuttujien X ja Y yhteisjakauman kertymäfunktio $F_{XY}(x, y)$.
 - (ii) Olkoon satunnaismuuttujan X reunajakauman kertymäfunktio $F_X(x)$.
Olkoon satunnaismuuttujan Y reunajakauman kertymäfunktio $F_Y(y)$.
- Määritelmä 2:
Satunnaismuuttujat X ja Y ovat **riippumattomia**, jos ja vain, jos

$$F_{XY}(x, y) = F_X(x)F_Y(y)$$

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Satunnaismuuttujien riippumattomuus:

Yleistys 1/2

- Oletukset:
 - (i) Olkoon satunnaismuuttujien X_i , $i = 1, 2, \dots, p$ yhteisjakauman pistetodennäköisyys- tai tiheysfunktio
$$f(x_1, x_2, \dots, x_p)$$
 - (ii) Olkoot satunnaismuuttujien X_i , $i = 1, 2, \dots, p$ reunajakaumien pistetodennäköisyys- tai tiheysfunktiot $f(x_i)$, $i = 1, 2, \dots, p$.

- Määritelmä 1:

Satunnaismuuttujat X_i , $i = 1, 2, \dots, p$ ovat **riippumattomia**, jos ja vain, jos

$$f(x_1, x_2, \dots, x_p) = f(x_1)f(x_2)\times \dots \times f(x_p)$$

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Satunnaismuuttujien riippumattomuus:

Yleistys 2/2

- Oletukset:
 - (i) Olkoon satunnaismuuttujien X_i , $i = 1, 2, \dots, p$ yhteisjakauman kertymäfunktio

$$F(x_1, x_2, \dots, x_p)$$

- (ii) Olkoot satunnaismuuttujien X_i , $i = 1, 2, \dots, p$ reunajakaumien kertymäfunktiot $F(x_i)$, $i = 1, 2, \dots, p$.

- Määritelmä 2:

Satunnaismuuttujat X_i , $i = 1, 2, \dots, p$ ovat **riippumattomia**, jos ja vain, jos

$$F(x_1, x_2, \dots, x_p) = F(x_1)F(x_2) \times \dots \times F(x_p)$$

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Satunnaismuuttujien riippumattomuus ja tapahtumien riippumattomuus

- Olkoot satunnaismuuttujat X ja Y *riippumattomia*.
- Tällöin

$$\Pr(a \leq X \leq b \text{ ja } c \leq Y \leq d) = \Pr(a \leq X \leq b) \Pr(c \leq Y \leq d)$$

- Huomautus:

Vrt. riippumattomien tapahtumien tulosääntö:

$$\Pr(A \cap B) = \Pr(A) \Pr(B)$$

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Satunnaismuuttujien riippumattomuus ja tapahtumien riippumattomuus: Perustelu

- Olkoot satunnaismuuttujat X ja Y jatkuvia ja *riippumattomia*.

$$\begin{aligned}\Pr(a \leq X \leq b \text{ ja } c \leq Y \leq d) &= \int_a^b \int_c^d f_{XY}(x, y) dy dx \\ &= \int_a^b \int_c^d f_X(x) f_Y(y) dy dx \\ &= \int_a^b f_X(x) \int_c^d f_Y(y) dy dx \\ &= \int_a^b f_X(x) \Pr(c \leq Y \leq d) dx \\ &= \Pr(c \leq Y \leq d) \int_a^b f_X(x) dx \\ &= \Pr(c \leq Y \leq d) \Pr(a \leq X \leq b)\end{aligned}$$

Satunnaismuuttujien riippumattomuus:

1. esimerkki nopanheitosta

- Heitetään *virheetöntä* noppaa kaksi kertaa:

X = tulos (silmäluku) 1. heitosta

Y = tulos (silmäluku) 2. heitosta

- Tarkastellaan satunnaismuuttujien X ja Y yhteisjakaumaa:

$$\Pr(X = x \text{ ja } Y = y) = \frac{1}{36}$$

$$= \frac{1}{6} \cdot \frac{1}{6}$$

$$= \Pr(X = x) \Pr(Y = y)$$

$$x = 1, 2, 3, 4, 5, 6; y = 1, 2, 3, 4, 5, 6$$

- Siten satunnaismuuttujat X ja Y ovat riippumattomia.

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Satunnaismuuttujien riippumattomuus:

2. esimerkki nopanheitosta

- Heitetään *virheetöntä* noppaa kaksi kertaa:

X = tulos (silmäluku) 1. heitosta

Y = tulos (silmäluku) 2. heitosta

$Z = X + Y$ = silmälukujen summa

- Tarkastellaan satunnaismuuttujien X ja Z yhteisjakaumaa.
- Esimerkiksi:

$$\Pr(X = 1 \text{ ja } Z = 8) = 0 \neq \frac{1}{6} \cdot \frac{5}{36} = \Pr(X = 1) \Pr(Z = 8)$$

- Siten satunnaismuuttujat X ja Z eivät ole riippumattomia.

Moniulotteiset satunnaismuuttujat ja todennäköisyysjakaumat

Kaksiulotteiset todennäköisyysjakaumat

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

>> Kaksiulotteisten jakaumien odotusarvot ja varianssit

Kovarianssi ja korrelaatio

Ehdolliset jakaumat ja odotusarvot

Diskreetin kaksiulotteisen satunnaismuuttujan funktion yleinen odotusarvo

- Olkoon satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio $f_{XY}(x, y)$.

- Olkoon

$$g : \mathbb{R}^2 \rightarrow$$

jatkuva funktio.

- Tällöin satunnaismuuttujan $g(X, Y)$ **odotusarvo** on *vakio*

$$E(g(X, Y)) = \sum_x \sum_y g(x, y) f_{XY}(x, y)$$

Jatkuvan kaksiulotteisen satunnaismuuttujan funktion yleinen odotusarvo

- Olkoon satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio $f_{XY}(x, y)$.

- Olkoon

$$g : \mathbb{R}^2 \rightarrow$$

jatkuva funktio.

- Tällöin satunnaismuuttujan $g(X, Y)$ **odotusarvo** on *vakio*

$$E(g(X, Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f_{XY}(x, y) dy dx$$

Reunajakaumien odotusarvot:

Diskreetit jakaumat 1/2

- Olkoon *diskreettien* satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio $f_{XY}(x, y)$.
- Satunnaismuuttujan X **odotusarvo** $E(X) = \mu_X$ yhtyy satunnaismuuttujan X *reunajakauman odotusarvoon*:

$$\begin{aligned} E(X) &= \sum_x \sum_y x f_{XY}(x, y) = \sum_x x \sum_y f_{XY}(x, y) \\ &= \sum_x x f_X(x) \end{aligned}$$

Reunajakaumien odotusarvot:

Diskreetit jakaumat 2/2

- Olkoon *diskreettien* satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio $f_{XY}(x, y)$.
- Satunnaismuuttujan Y **odotusarvo** $E(Y) = \mu_Y$ yhtyy satunnaismuuttujan Y *reunajakauman odotusarvoon*:

$$\begin{aligned} E(Y) &= \sum_x \sum_y y f_{XY}(x, y) = \sum_y y \sum_x f_{XY}(x, y) \\ &= \sum_y y f_Y(y) \end{aligned}$$

Reunajakaumien odotusarvot:

Jatkuvat jakaumat 1/2

- Olkoon *jatkuvien* satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio $f_{XY}(x, y)$.
- Satunnaismuuttujan X **odotusarvo** $E(X) = \mu_X$ yhtyy satunnaismuuttujan X *reunajakauman odotusarvoon*:

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_{XY}(x, y) dy dx = \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f_{XY}(x, y) dy dx \\ &= \int_{-\infty}^{+\infty} x f_X(x) dx \end{aligned}$$

Reunajakaumien odotusarvot:

Jatkuvat jakaumat 2/2

- Olkoon *jatkuvien* satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio $f_{XY}(x, y)$.
- Satunnaismuuttujan Y **odotusarvo** $E(Y) = \mu_Y$ yhtyy satunnaismuuttujan Y *reunajakauman odotusarvoon*:

$$\begin{aligned} E(Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f_{XY}(x, y) dy dx = \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f_{XY}(x, y) dx dy \\ &= \int_{-\infty}^{+\infty} y f_Y(y) dy \end{aligned}$$

Odotusarvot ja todennäköisyysmassan painopiste

- Olkoot satunnaismuuttujien X ja Y odotusarvot

$$E(X) = \mu_X \quad E(Y) = \mu_Y$$

- Tällöin kaksiulotteisen satunnaismuuttujan (X, Y) odotusarvo on järjestetty pari

$$(E(X), E(Y)) = (\mu_X, \mu_Y)$$

- Satunnaismuuttujien X ja Y odotusarvojen muodostama järjestetty pari (μ_X, μ_Y) määrää satunnaismuuttujien X ja Y yhteisjakauman todennäköisyysmassan **painopisteen**.

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Summan ja erotuksen odotusarvot

- Satunnaismuuttujien X ja Y **summan $X + Y$ odotusarvo:**

$$E(X + Y) = E(X) + E(Y)$$

- Satunnaismuuttujien X ja Y **erotuksen $X - Y$ odotusarvo:**

$$E(X - Y) = E(X) - E(Y)$$

- **Huomautus:**

Satunnaismuuttujien summan ja erotuksen odotusarvojen kaavat on esitetty ilman perustelua luvussa **Jakaumien tunnusluvut**.

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Summan ja erotuksen odotusarvot:

Perustelu 1/2

- Esitetään satunnaismuuttujien summan odotusarvoa koskevan tuloksen perustelu kahden *jatkuvan* satunnaismuuttujan tapauksessa.
- Olkoot satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio

$$f_{XY}(x, y)$$

ja X :n ja Y :n reunajakaumien tiheysfunktiot vastaavasti

$$f_X(x) \text{ ja } f_Y(y)$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Summan ja erotuksen odotusarvot:

Perustelu 2/2

- Tällöin $E(X \pm Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \pm y) f_{XY}(x, y) dy dx$
 $= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [x f_{XY}(x, y) \pm y f_{XY}(x, y)] dy dx$
 $= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_{XY}(x, y) dy dx \pm \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f_{XY}(x, y) dy dx$
 $= \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f_{XY}(x, y) dy dx \pm \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f_{XY}(x, y) dx dy$
 $= \int_{-\infty}^{+\infty} x f_X(x) dx \pm \int_{-\infty}^{+\infty} y f_Y(y) dy$
 $= E(X) \pm E(Y)$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Lineaarikombinaation odotusarvo

- Olkoot satunnaismuuttujien X_i , $i = 1, 2, \dots, k$ *odotusarvot*

$$E(X_i) = \mu_i, i = 1, 2, \dots, k$$

ja olkoot a_i , $i = 1, 2, \dots, k$ *vakioita*.

- Tällöin satunnaismuuttujien X_i , $i = 1, 2, \dots, k$ **lineaarikombinaation**

$$Y = a_1 X_1 + a_2 X_2 + \dots + a_k X_k$$

odotusarvo on

$$\begin{aligned} E(Y) &= E(a_1 X_1 + a_2 X_2 + \dots + a_k X_k) \\ &= a_1 E(X_1) + a_2 E(X_2) + \dots + a_k E(X_k) \\ &= a_1 \mu_1 + a_2 \mu_2 + \dots + a_k \mu_k \end{aligned}$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Riippumattomuus ja tulon odotusarvo

- Olkoot satunnaismuuttujien X ja Y *odotusarvot*

$$E(X) = \mu_X \quad E(Y) = \mu_Y$$

- Jos satunnaismuuttujat X ja Y ovat *riippumattomia*, niin **tulon XY odotusarvo on odotusarvojen tulo:**

$$\begin{aligned} E(XY) &= E(X)E(Y) \\ &= \mu_X \mu_Y \end{aligned}$$

- Huomautus:

Käänteinen *ei päde*: Siitä, että

$$E(XY) = E(X)E(Y)$$

ei seuraa, että satunnaismuuttujat X ja Y ovat riippumattomia.

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Riippumattomuus ja tulon odotusarvo:

Perustelu 1/2

- Esitetään riippumattomien satunnaismuuttujien tulon odotusarvoa koskevan tuloksen perustelu kahden *jatkuvan* satunnaismuuttujan tapauksessa.

- Olkoot satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio

$$f_{XY}(x, y)$$

ja X :n ja Y :n reunajakaumien tiheysfunktiot vastaavasti

$$f_X(x) \text{ ja } f_Y(y)$$

- Olkoot satunnaismuuttujat X ja Y riippumattomia, jolloin

$$f_{XY}(x, y) = f_X(x) f_Y(y)$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Riippumattomuus ja tulon odotusarvo:

Perustelu 2/2

- Tällöin $E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf_{XY}(x, y)dydx$
 $= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xyf_X(x)f_Y(y)dydx$
 $= \int_{-\infty}^{+\infty} xf_X(x) \int_{-\infty}^{+\infty} yf_Y(y)dydx$
 $= \int_{-\infty}^{+\infty} xf_X(x)E(Y)dx$
 $= E(Y) \int_{-\infty}^{+\infty} xf_X(x)dx$
 $= E(Y)E(X)$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Riippumattomuus ja tulon odotusarvo:

Yleistys

- Olkoot satunnaismuuttujien X_i , $i = 1, 2, \dots, k$ odotusarvot

$$E(X_i) = \mu_i, i = 1, 2, \dots, k$$

- Jos satunnaismuuttujat X_i , $i = 1, 2, \dots, k$ ovat riippumattomia, niin

$$\begin{aligned} E(X_1 X_2 \dots X_k) &= E(X_1) E(X_2) \dots E(X_k) \\ &= \mu_1 \mu_2 \dots \mu_k \end{aligned}$$

- Huomautus:

Käänteinen *ei päde*: Siitä, että

$$E(X_1 X_2 \dots X_k) = E(X_1) E(X_2) \dots E(X_k)$$

ei seuraa, että satunnaismuuttujat X_i , $i = 1, 2, \dots, k$ ovat riippumattomia.

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Reunajakaumien varianssit

- Olkoot satunnaismuuttujien X ja Y odotusarvot

$$E(X) = \mu_X \quad E(Y) = \mu_Y$$

- Satunnaismuuttujien X ja Y **varianssit** yhtyvät vastaavien *reunajakaumien variansseihin*:

$$\begin{aligned} \text{Var}(X) &= D^2(X) = \sigma_X^2 \\ &= E[(X - \mu_X)^2] \end{aligned}$$

$$\begin{aligned} \text{Var}(Y) &= D^2(Y) = \sigma_Y^2 \\ &= E[(Y - \mu_Y)^2] \end{aligned}$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Varianssi vaihtelun mittana

- Satunnaismuuttujan X varianssi

$$\text{Var}(X)$$

kuvaa satunnaismuuttujan X todennäköisyysjakauman *todennäköisyysmassan vaihtelua* satunnaismuuttujan X oman odotusarvon $E(X)$ ympärillä.

- Satunnaismuuttujan Y varianssi

$$\text{Var}(Y)$$

kuvaa satunnaismuuttujan Y todennäköisyysjakauman *todennäköisyysmassan vaihtelua* satunnaismuuttujan Y oman odotusarvon $E(Y)$ ympärillä.

Reunajakaumien varianssit:

Diskreetit jakaumat

- Olkoon *diskreettien* satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio $f_{XY}(x, y)$ ja vastaavien reunajakaumien pistetodennäköisyysfunktiot $f_X(x)$ ja $f_Y(y)$.
- Tällöin satunnaismuuttujien X ja Y **varianssit** ovat *vakioita*

$$D^2(X) = \sum_x (x - \mu_X)^2 f_X(x)$$

$$D^2(Y) = \sum_y (y - \mu_Y)^2 f_Y(y)$$

jossa

$$\mu_X = E(X)$$

$$\mu_Y = E(Y)$$

Reunajakaumien varianssit:

Jatkuvat jakaumat

- Olkoon *jatkuvien* satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio $f_{XY}(x, y)$ ja vastaavien reunajakaumien tiheysfunktiot $f_X(x)$ ja $f_Y(y)$.
- Tällöin satunnaismuuttujien X ja Y **varianssit** ovat *vakioita*

$$D^2(X) = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f_X(x) dx$$

$$D^2(Y) = \int_{-\infty}^{+\infty} (y - \mu_Y)^2 f_Y(y) dy$$

jossa

$$\mu_X = E(X)$$

$$\mu_Y = E(Y)$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Varianssien vaihtoehtoiset laskukaavat

- Satunnaismuuttujien X ja Y varianssien kaavat voidaan kirjoittaa seuraaviin yhtäpitäviin muotoihin:

$$\begin{aligned}D^2(X) &= E[(X - \mu_X)^2] \\ &= E(X^2) - \mu_X^2 \\ &= E(X^2) - [E(X)]^2\end{aligned}$$

$$\begin{aligned}D^2(Y) &= E[(Y - \mu_Y)^2] \\ &= E(Y^2) - \mu_Y^2 \\ &= E(Y^2) - [E(Y)]^2\end{aligned}$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Varianssien vaihtoehtoiset laskukaavat:

Perustelu

- Esitetään varianssin vaihtoehtoisen laskukaavan perustelu satunnaismuuttujalle X :

$$\begin{aligned} D^2(X) &= E[(X - \mu_X)^2] \\ &= E[X^2 - 2\mu_X X + \mu_X^2] \\ &= E(X^2) - E(2\mu_X X) + E(\mu_X^2) \\ &= E(X^2) - 2\mu_X E(X) + \mu_X^2 \\ &= E(X^2) - 2\mu_X \mu_X + \mu_X^2 \\ &= E(X^2) - \mu_X^2 \end{aligned}$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Reunajakaumien standardipoikkeamat

- Olkoot satunnaismuuttujien X ja Y *odotusarvot*

$$E(X) = \mu_X \quad E(Y) = \mu_Y$$

- Satunnaismuuttujien X ja Y **standardipoikkeamat** yhtyvät vastaavien *reunajakaumien standardipoikkeamiin*:

$$\begin{aligned} D(X) &= \sigma_X \\ &= \sqrt{E[(X - \mu_X)^2]} \end{aligned}$$

$$\begin{aligned} D(Y) &= \sigma_Y \\ &= \sqrt{E[(Y - \mu_Y)^2]} \end{aligned}$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Standardipoikkeama vaihtelun mittana

- Satunnaismuuttujan X standardipoikkeama

$$D(X)$$

kuvaa satunnaismuuttujan X todennäköisyysjakauman *todennäköisyysmassan vaihtelua* satunnaismuuttujan X oman odotusarvon $E(X)$ ympärillä.

- Satunnaismuuttujan Y standardipoikkeama

$$D(Y)$$

kuvaa satunnaismuuttujan Y todennäköisyysjakauman *todennäköisyysmassan vaihtelua* satunnaismuuttujan Y oman odotusarvon $E(Y)$ ympärillä.

Odotusarvo, varianssi ja standardipoikkeama: Suureiden dimensio eli laatu

- Satunnaismuuttujalla sekä sen *odotusarvolla* ja *standardipoikkeamalla on aina sama dimensio eli laatu.*

Esimerkki:

Jos satunnaismuuttujan laatuna on m (metri), niin myös niiden odotusarvon ja standardipoikkeaman laatuna on m.

- Satunnaismuuttujan ja sen *varianssilla ei ole sama dimensio eli laatu.*

Esimerkki:

Jos satunnaismuuttujan laatuna on m (metri), niin sen varianssin laatuna on m^2 .

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Odotusarvot ja standardipoikkeamat:

1. esimerkki nopanheitosta

- Heitetään *virheetöntä* noppaa kaksi kertaa:

X = tulos (silmäluku) 1. heitosta

Y = tulos (silmäluku) 2. heitosta

- Tarkastellaan satunnaismuuttujien X ja Y yhteisjakaumaa.
- *Odotusarvot, varianssit ja standardipoikkeamat:*

$$E(X) = \sum_{x=1}^6 x \Pr(X = x) = \frac{1}{6} \sum_{x=1}^6 x = \frac{21}{6} = 3.5 = E(Y)$$

$$E(X^2) = \sum_{x=1}^6 x^2 \Pr(X = x) = \frac{1}{6} \sum_{x=1}^6 x^2 = \frac{91}{6} = E(Y^2)$$

$$D^2(X) = E(X^2) - [E(X)]^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{35}{12} = 2.917 = D^2(Y)$$

$$D(X) = \sqrt{2.917} = 1.708 = D(Y)$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Odotusarvot ja standardipoikkeamat:

2. esimerkki nopanheitosta 1/2

- Heitetään *virheetöntä* noppaa kaksi kertaa:

X = tulos (silmäluku) 1. heitosta

Y = tulos (silmäluku) 2. heitosta

$Z = X + Y$ = silmälukujen summa

- Tarkastellaan satunnaismuuttujien X ja Z yhteisjakaumaa.
- *Odotusarvot ja 2. momentit:*

$$E(X) = \sum_{x=1}^6 x \Pr(X = x) = \frac{21}{6} = 3.5 = E(Y)$$

$$E(Z) = \sum_{z=1}^{12} z \Pr(Z = z) = \frac{252}{36} = 7$$

$$E(X^2) = \sum_{x=1}^6 x^2 \Pr(X = x) = \frac{91}{6} = E(X^2)$$

$$E(Z^2) = \sum_{z=1}^{12} z^2 \Pr(Z = z) = \frac{1974}{36}$$

Kaksiulotteisten jakaumien odotusarvot ja varianssit
Odotusarvot ja standardipoikkeamat:
2. esimerkki nopanheitosta 2/2

- *Varianssit ja standardipoikkeamat :*

$$D^2(X) = E(X^2) - [E(X)]^2 = \frac{35}{12} = 2.917 = D^2(Y)$$

$$D(X) = \sqrt{2.917} = 1.708 = D(Y)$$

$$D^2(Z) = E(Z^2) - [E(Z)]^2 = \frac{210}{36} = 5.833$$

$$D(Z) = \sqrt{5.833} = 2.415$$

Moniulotteiset satunnaismuuttajat ja todennäköisyysjakaumat

Kaksiulotteiset todennäköisyysjakaumat

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Kaksiulotteisten jakaumien odotusarvot ja varianssit

>> Kovarianssi ja korrelaatio

Ehdolliset jakaumat ja odotusarvot

Kovarianssi ja korrelaatio

Kovarianssi

- Olkoot satunnaismuuttujien X ja Y *odotusarvot*

$$E(X) = \mu_X \quad E(Y) = \mu_Y$$

- Satunnaismuuttujien X ja Y **kovarianssi** on *vakio*

$$\begin{aligned} \text{Cov}(X, Y) &= \sigma_{XY} \\ &= E[(X - \mu_X)(Y - \mu_Y)] \end{aligned}$$

- Satunnaismuuttujien X ja Y kovarianssi

$$\text{Cov}(X, Y)$$

kuvaa satunnaismuuttujien X ja Y yhteisjakauman todennäköisyysmassan **yhteisvaihtelua** satunnaismuuttujien X ja Y odotusarvojen $E(X)$ ja $E(Y)$ määräämän pisteen $(E(X), E(Y))$ ympärillä.

Kovarianssi: Diskreetit jakaumat

- Olkoon *diskreettien* satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyysfunktio $f_{XY}(x, y)$.
- Tällöin satunnaismuuttujien X ja Y **kovarianssi** on *vakio*

$$\text{Cov}(X, Y) = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f_{XY}(x, y)$$

jossa

$$\mu_X = E(X)$$

$$\mu_Y = E(Y)$$

Kovarianssi: Jatkuvat jakaumat

- Olkoon *jatkuvien* satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio $f_{XY}(x, y)$.
- Tällöin satunnaismuuttujien X ja Y **kovarianssi** on *vakio*

$$\text{Cov}(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)(y - \mu_Y) f_{XY}(x, y) dx dy$$

jossa

$$\mu_X = E(X)$$

$$\mu_Y = E(Y)$$

Kovarianssi:

Vaihtoehtoinen laskukaava

- Satunnaismuuttujien X ja Y kovarianssin kaava voidaan kirjoittaa seuraaviin yhtäpitäviin muotoihin:

$$\begin{aligned}\text{Cov}(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E(XY) - \mu_X \mu_Y \\ &= E(XY) - E(X)E(Y)\end{aligned}$$

- Huomautus:

Vrt. kovarianssin vaihtoehtoista laskukaavaa varianssin vaihtoehtoiseen laskukaavaan.

Kovarianssin vaihtoehtoinen laskukaava: Perustelu

- Esitetään satunnaismuuttujien kovarianssin vaihtoehtoisen laskukaavan perustelu:

$$\begin{aligned}\text{Cov}(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y] \\ &= E(XY) - E(\mu_X Y) - E(\mu_Y X) + E(\mu_X \mu_Y) \\ &= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X \mu_Y \\ &= E(XY) - \mu_X \mu_Y - \mu_Y \mu_X + \mu_X \mu_Y \\ &= E(XY) - \mu_X \mu_Y\end{aligned}$$

Kovarianssi ja varianssit

- Satunnaismuuttujien kovarianssit itsensä kanssa yhtyvät satunnaismuuttujien variansseihin:

$$\text{Cov}(X, X) = E[(X - \mu_X)(X - \mu_X)]$$

$$= \text{Var}(X)$$

$$= \sigma_X^2$$

$$\text{Cov}(Y, Y) = E[(Y - \mu_Y)(Y - \mu_Y)]$$

$$= \text{Var}(Y)$$

$$= \sigma_Y^2$$

Odotusarvot, varianssit ja kovarianssi sekä lineaarimuunnokset 1/2

- Olkoot satunnaismuuttujien X ja Y odotusarvot, varianssit ja kovarianssi:

$$E(X) = \mu_X \quad \text{Var}(X) = \sigma_X^2$$

$$E(Y) = \mu_Y \quad \text{Var}(Y) = \sigma_Y^2$$

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = \sigma_{XY}$$

- Olkoot

$$W = a + bX$$

$$Z = c + dY$$

jossa $a, b, c, d \in \mathbb{P}$ ovat reaalisia vakioita.

Odotusarvot, varianssit ja kovarianssi sekä lineaarimuunnokset 2/2

- Tällöin pätevät seuraavat kaavat:

$$E(W) = a + b E(X) = a + b\mu_X$$

$$\text{Var}(W) = b^2 \text{Var}(X) = b^2 \sigma_X^2$$

$$E(Z) = c + d E(Y) = c + d\mu_Y$$

$$\text{Var}(Z) = d^2 \text{Var}(Y) = d^2 \sigma_Y^2$$

$$\text{Cov}(W, Z) = bd \text{Cov}(X, Y) = bd \sigma_{XY}$$

Kovarianssi ja lineaarimuunnokset: Perustelu

- Jos satunnaismuuttujien X ja Y kovarianssi on $\text{Cov}(X, Y)$, niin

$$\begin{aligned}\text{Cov}(W, Z) &= E[(W - E(W))(Z - E(Z))] \\ &= E[(a + bX - E(a + bX))(c + dY - E(c + dY))] \\ &= E[(bX - E(bX))(dY - E(dY))] \\ &= bd E[(X - E(X))(Y - E(Y))] \\ &= bd \text{Cov}(X, Y)\end{aligned}$$

Summan ja erotuksen varianssit

- Olkoot satunnaismuuttujien X ja Y varianssit

$$\text{Var}(X) = \sigma_X^2 \quad \text{Var}(Y) = \sigma_Y^2$$

ja kovarianssi

$$\text{Cov}(X, Y) = \sigma_{XY}$$

- Tällöin

$$\begin{aligned} \text{Var}(X \pm Y) &= \text{Var}(X) + \text{Var}(Y) \pm 2\text{Cov}(X, Y) \\ &= \sigma_X^2 + \sigma_Y^2 \pm 2\sigma_{XY} \end{aligned}$$

Summan ja erotuksen varianssit: Perustelu

- Summan ja erotuksen varianssia koskevat kaavat nähdään oikeaksi seuraavalla tavalla:

$$\begin{aligned}\text{Var}(X \pm Y) &= E[X \pm Y - E(X \pm Y)]^2 \\ &= E[X \pm Y - E(X) \mp E(Y)]^2 \\ &= E[(X - E(X)) \pm (Y - E(Y))]^2 \\ &= E[(X - E(X))^2 + (Y - E(Y))^2 \\ &\quad \pm 2(X - E(X))(Y - E(Y))] \\ &= E[(X - E(X))^2] + E[(Y - E(Y))^2] \\ &\quad \pm 2E[(X - E(X))(Y - E(Y))] \\ &= \text{Var}(X) + \text{Var}(Y) \pm 2\text{Cov}(X, Y)\end{aligned}$$

Riippumattomuus ja kovarianssi

- Jos satunnaismuuttujat X ja Y ovat *riippumattomia*, niin

$$\text{Cov}(X, Y) = 0$$

- Huomautus:

Käänteinen *ei päde*: Siitä, että

$$\text{Cov}(X, Y) = 0$$

ei välttämättä seuraa, että satunnaismuuttujat X ja Y olisivat riippumattomia.

Riippumattomuus ja kovarianssi: Perustelu

- Jos satunnaismuuttujat X ja Y ovat *riippumattomia*, niin

$$E(XY) = E(X)E(Y)$$

- Siten

$$\begin{aligned}\text{Cov}(X, Y) &= E(XY) - E(X)E(Y) \\ &= E(X)E(Y) - E(X)E(Y) \\ &= 0\end{aligned}$$

Riippumattomuus sekä summan ja erotuksen varianssit

- Olkoot satunnaismuuttujien X ja Y varianssit

$$\text{Var}(X) = \sigma_X^2 \quad \text{Var}(Y) = \sigma_Y^2$$

- Olkoot X ja Y ovat *riippumattomia*, jolloin

$$\text{Cov}(X, Y) = 0$$

- Tällöin

$$\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$$

- Huomautus:

Riippumattomien satunnaismuuttujien summan ja erotuksen varianssien kaavat on esitetty ilman perustelua luvussa **Jakaumien tunnusluvut**.

Korrelaatiokerroin 1/2

- Oletetaan, että satunnaismuuttujilla X ja Y on seuraavat odotusarvot, varianssit ja kovarianssi:

$$E(X) = \mu_X \quad \text{Var}(X) = D^2(X) = \sigma_X^2$$

$$E(Y) = \mu_Y \quad \text{Var}(Y) = D^2(Y) = \sigma_Y^2$$

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = \sigma_{XY}$$

Korrelaatiokerroin 2/2

- Satunnaismuuttujien X ja Y **korrelaatiokerroin** on *vakio*

$$\begin{aligned}\text{Cor}(X, Y) &= \rho_{XY} \\ &= \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} \\ &= \frac{\text{Cov}(X, Y)}{D(X)D(Y)} \\ &= \frac{\sigma_{XY}}{\sigma_X \sigma_Y}\end{aligned}$$

Korrelaatiokerroin

lineaarisen riippuvuuden mittana

- Satunnaismuuttujien X ja Y korrelaatiokerroin

$$\text{Cor}(X, Y)$$

mittaa satunnaismuuttujien X ja Y **lineaarisen riippuvuuden** *voimakkuutta*:

- (i) Mitä *suurempi* on

$$|\text{Cor}(X, Y)|$$

sitä *voimakkaampaa* on satunnaismuuttujien X ja Y välinen lineaarinen riippuvuus.

- (ii) Mitä *pienempi* on

$$|\text{Cor}(X, Y)|$$

sitä *heikompa* on satunnaismuuttujien X ja Y välinen lineaarinen riippuvuus.

Kovarianssi ja korrelaatio: Suureiden dimensio eli laatu

- Huomaa, että satunnaismuuttujien korrelaatio *on dimensioton eli laaduton suure* toisin kuin niiden kovarianssi.

Korrelaatiokerroin ja lineaarimuunnokset

- Olkoon satunnaismuuttujien X ja Y korrelaatiokerroin $\text{Cor}(X, Y)$.

- Olkoot

$$W = a + bX$$

$$Z = c + dY$$

- Tällöin

$$\text{Cor}(W, Z) = \text{sgn}(bd) \text{Cor}(X, Y)$$

jossa sgn on ns. *merkkifunktio*:

$$\text{sgn}(x) = \begin{cases} +1, & \text{jos } x > 0 \\ 0, & \text{jos } x = 0 \\ -1, & \text{jos } x < 0 \end{cases}$$

Lineaarimuunnosten korrelaatiokerroin: Perustelu

- Olkoon satunnaismuuttujien X ja Y korrelaatiokerroin $\text{Cor}(X, Y)$.
- Olkoot

$$W = a + bX$$

$$Z = c + dY$$

- Tällöin

$$\begin{aligned}\text{Cor}(W, Z) &= \frac{\text{Cov}(W, Z)}{\sqrt{\text{Var}(W) \text{Var}(Z)}} \\ &= \frac{bd \text{Cov}(X, Y)}{\sqrt{b^2 \text{Var}(X) d^2 \text{Var}(Y)}} \\ &= \text{sgn}(bd) \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} \\ &= \text{sgn}(bd) \text{Cor}(X, Y)\end{aligned}$$

Korrelaatiokerroin: Ominaisuudet

- Olkoon satunnaismuuttujien X ja Y korrelaatiokerroin $\text{Cor}(X, Y)$.
- Tällöin
 - (i) $-1 \leq \text{Cor}(X, Y) \leq +1$
 - (ii) Jos X ja Y ovat *riippumattomia*, niin $\text{Cor}(X, Y) = 0$
 - (iii) $\text{Cor}(X, Y) = \pm 1$, jos ja vain, jos
$$Y = \alpha + \beta X,$$
jossa α ja β ovat reaalisia vakioita, $\beta \neq 0$

Kovarianssi ja korrelaatio

Korrelaatiokerroin:

Ominaisuuden (i) perustelu 1/2

- Väite:

$$-1 \leq \text{Cor}(X, Y) \leq +1$$

- Perustelu:

(1) Olkoot W ja Z satunnaismuuttujia ja olkoon $a \in \mathbb{R}$ vakio.

(2) $W^2 \geq 0, Z^2 \geq 0 \Rightarrow E(W^2) \geq 0, E(Z^2) \geq 0$

(3) $(aW - Z)^2 \geq 0 \Rightarrow E(aW - Z)^2 \geq 0$

(4) Kohdasta (3) seuraa, että

$$E(aW - Z)^2 = a^2E(W^2) - 2aE(WZ) + E(Z^2) \geq 0$$

(5) Valitaan:

$$a = \frac{E(WZ)}{E(W^2)}$$

(6) Siten

$$E(aW - Z)^2 = -\frac{[E(WZ)]^2}{E(W^2)} + E(Z^2) \geq 0$$

josta seuraa, että

$$\frac{[E(WZ)]^2}{E(W^2)E(Z^2)} \leq 1 \quad (*)$$

(7) Väite seuraa epäyhtälöstä (*), kun valitaan

$$W = X - E(X) \text{ ja } Z = Y - E(Y)$$

ja otetaan saadusta epäyhtälöstä neliöjuuri.

- Huomautus:

Epäyhtälöstä (*) seuraa eräs ns. *Schwarzin epäyhtälön* monista muodoista:

$$[E(WZ)]^2 \leq E(W^2)E(Z^2)$$

Korrelaatiokerroin: Ominaisuuden (ii) perustelu

- Väite:

Jos X ja Y ovat *riippumattomia*, niin

$$\text{Cor}(X, Y) = 0$$

- Perustelu:

Väite seuraa siitä, että jos X ja Y ovat riippumattomia, niin

$$\text{Cov}(X, Y) = 0$$

- Huomautus:

Käänteinen *ei päde*: Siitä, että

$$\text{Cov}(X, Y) = \text{Cor}(X, Y) = 0$$

ei välttämättä seuraa, että satunnaismuuttujat X ja Y olisivat riippumattomia.

Korrelaatiokerroin:

Ominaisuuden (iii) perustelu 1/3

- Väite:

$$\text{Cor}(X, Y) = \pm 1 \Rightarrow Y = \alpha + \beta X, \alpha \in \mathbb{R}, \beta \in \mathbb{R}, \beta \neq 0$$

- Perustelu:

- (1) Ominaisuuden (i) perustelussa kohdan (4) epäyhtälön vasen puoli on 2. asteen polynomi muuttujan a suhteen:

$$h(a) = E(aW - Z)^2 = a^2 E(W^2) - 2a E(WZ) + E(Z^2)$$

- (2) Ominaisuuden (i) perustelun kohdan (3) mukaan $h(a) \geq 0 \forall a$.

- (3) $h(a) > 0$ kaikille a , jos ja vain jos 2. asteen yhtälön $h(a) = 0$ *diskriminantti*

$$D = 4[E(WZ)]^2 - 4E(W^2)E(Z^2) < 0$$

- Huomautus:

Korrelaatiokerroimen ominaisuus (i) seuraa myös kohdan (3) epäyhtälöstä.

Korrelaatiokerroin:

Ominaisuuden (iii) perustelu 2/3

- (4) $h(a) = 0$ täsmälleen silloin, kun 2. asteen yhtälön $h(a) = 0$ diskriminantti

$$D = 4[E(WZ)]^2 - 4E(W^2)E(Z^2) = 0$$

- (5) Siten $\text{Cor}(X, Y) = \pm 1$ täsmälleen silloin, kun

$$E(aW - Z)^2 = 0$$

mikä on yhtäpitävää sen kanssa, että

$$\Pr(aW - Z = 0) = 1$$

- (6) Siten

$$Z = aW$$

todennäköisyydellä 1.

- (7) Väite seuraa, kun valitaan $W = X - E(X)$ ja $Z = Y - E(Y)$ ja merkitään

$$\beta = a \text{ ja } \alpha = E(Y) - a E(X)$$

Korrelaatiokerroin:

Ominaisuuden (iii) perustelu 3/3

- Väite:

$$\text{Cor}(X, Y) = \pm 1 \iff Y = \alpha + \beta X, \alpha \in \mathbb{R}, \beta \in \mathbb{R}, \beta \neq 0$$

- Perustelu:

Koska $Y = \alpha + \beta X$,

$$\begin{aligned}\text{Cor}(X, Y) &= \text{Cor}(X, \alpha + \beta X) \\ &= \text{sgn}(\beta) \text{Cor}(X, X) \\ &= \text{sgn}(\beta) \\ &= \pm 1\end{aligned}$$

jossa $\text{sgn}(\cdot)$ on ns. *merkkifunktio*:

$$\text{sgn}(x) = \begin{cases} +1, & \text{jos } x > 0 \\ 0, & \text{jos } x = 0 \\ -1, & \text{jos } x < 0 \end{cases}$$

Korrelaatiokerroin:

1. esimerkki nopanheitosta

- Heitetään *virheetöntä* noppaa kaksi kertaa:

$X =$ tulos (silmäluku) 1. heitosta

$Y =$ tulos (silmäluku) 2. heitosta

- Tarkastellaan satunnaismuuttujien X ja Y yhteisjakaumaa.
- *Odotusarvot, varianssit ja standardipoikkeamat* (laskettu aikaisemmin):

$$E(X) = E(Y) = 21/6 = 3.5$$

$$D^2(X) = D^2(Y) = 35/12 = 2.917$$

$$D(X) = D(Y) = 1.708$$

- Koska satunnaismuuttujat X ja Y ovat *riippumattomia*, niin
$$\text{Cor}(X, Y) = \text{Cov}(X, Y) = 0$$

Korrelaatiokerroin:

2. esimerkki nopanheitosta 1/2

- Heitetään *virheetöntä* noppaa kaksi kertaa:

X = tulos (silmäluku) 1. heitosta

Y = tulos (silmäluku) 2. heitosta

$Z = X + Y$ = silmälukujen summa

- Tarkastellaan satunnaismuuttujien X ja Z yhteisjakaumaa.
- *Odotusarvot, varianssit ja standardipoikkeamat* (laskettu aikaisemmin):

$$E(X) = 21/6 = 3.5$$

$$E(Z) = 252/36 = 7$$

$$D^2(X) = 35/12 = 2.917$$

$$D^2(Z) = 210/36 = 5.833$$

$$D(X) = 1.708$$

$$D(Z) = 2.415$$

Kovarianssi ja korrelaatio

Korrelaatiokerroin:

2. esimerkki nopanheitosta 2/2

- Satunnaismuuttujat X ja Z eivät ole riippumattomia.
- Lasketaan ensin *kovarianssi*:

$$E(XZ) = \sum_{x=1}^6 \sum_{z=2}^{12} xz \Pr(X = x, Z = z) = \frac{987}{36}$$

$$\text{Cov}(X, Z) = E(XZ) - E(X)E(Z) = \frac{987}{36} - \frac{21}{6} \cdot \frac{42}{6} = \frac{105}{36} = 2.917$$

- *Korrelaatiokertoimen* arvo on:

$$\text{Cor}(X, Z) = \frac{\text{Cov}(X, Z)}{D(X)D(Z)} = \frac{1}{\sqrt{2}} = 0.707$$

Kovarianssi ja korrelaatio

Korreloimattomuus

- Jos

$$\text{Cov}(X, Y) = 0$$

ja yhtäpitävästi

$$\text{Cor}(X, Y) = 0$$

niin sanomme, että satunnaismuuttujat X ja Y ovat **korreloimattomia**.

- Huomautus:

Satunnaismuuttujien korreloimattomuudesta *ei välttämättä seuraa* niiden riippumattomuus.

Moniulotteiset satunnaismuuttajat ja todennäköisyysjakaumat

Kaksiulotteiset todennäköisyysjakaumat

Kaksiulotteisten jakaumien reunajakaumat ja riippumattomuus

Kaksiulotteisten jakaumien odotusarvot ja varianssit

Kovarianssi ja korrelaatio

>> Ehdolliset jakaumat ja odotusarvot

Ehdolliset jakaumat ja odotusarvot

Ehdollinen todennäköisyys

- Olkoot A ja B tapahtumia ja $\Pr(B) \neq 0$.
- Tapahtuman A *ehdollinen todennäköisyys* tapahtuman B suhteen on

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

- *Ehdollisen jakauman määritelmä* mukailee ehdollisen todennäköisyyden määritelmää.

Ehdolliset jakaumat

- Olkoon satunnaismuuttujien X ja Y yhteisjakauman *pistetodennäköisyys-* tai *tiheysfunktio* $f_{XY}(x, y)$.
- Olkoot satunnaismuuttujien X ja Y *reunajakaumien* pistetodennäköisyys- tai tiheysfunktiot $f_X(x)$ ja $f_Y(y)$.
- Satunnaismuuttujan X **ehdollinen jakauma** satunnaismuuttujan Y suhteen (ehdolla $Y = y$) on

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)}, \text{ jos } f_Y(y) > 0$$

- Satunnaismuuttujan Y **ehdollinen jakauma** satunnaismuuttujan X suhteen (ehdolla $X = x$) on

$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)}, \text{ jos } f_X(x) > 0$$

Ehdolliset jakaumat: Kommentteja

- Satunnaismuuttujan X ehdollinen jakauma

$$f_{X|Y}(x|y)$$

satunnaismuuttujan Y suhteen (ehdolla $Y = y$) *riippuu yleisessä tapauksessa ehtomuuttujan Y arvosta y .*

- Satunnaismuuttujan Y ehdollinen jakauma

$$f_{Y|X}(y|x)$$

satunnaismuuttujan X suhteen (ehdolla $X = x$) *riippuu yleisessä tapauksessa ehtomuuttujan X arvosta x .*

Riippumattomuus ja ehdolliset jakaumat 1/3

- Olkoon satunnaismuuttujien X ja Y yhteisjakauman *pistetodennäköisyys-* tai *tiheysfunktio* $f_{XY}(x, y)$.
- Olkoot satunnaismuuttujien X ja Y *reunajakaumien* *pistetodennäköisyys-* tai *tiheysfunktiot* $f_X(x)$ ja $f_Y(y)$.
- Olkoot satunnaismuuttujat X ja Y *riippumattomia*.
- Tällöin

$$f_{XY}(x, y) = f_X(x) f_Y(y)$$

Riippumattomuus ja ehdolliset jakaumat 2/3

- Jos satunnaismuuttujat X ja Y ovat riippumattomia, satunnaismuuttujan X *ehdollinen jakauma* satunnaismuuttujan Y suhteen yhtyy satunnaismuuttujan X *reunajakaumaan*:

$$f_{X|Y}(x|y) = f_X(x) , \text{ jos } f_Y(y) > 0$$

- Jos satunnaismuuttujat X ja Y ovat riippumattomia, satunnaismuuttujan Y *ehdollinen jakauma* satunnaismuuttujan X suhteen yhtyy satunnaismuuttujan Y *reunajakaumaan*:

$$f_{Y|X}(y|x) = f_Y(y) , \text{ jos } f_X(x) > 0$$

Riippumattomuus ja ehdolliset jakaumat 3/3

- Olkoot satunnaismuuttujat X ja Y riippumattomia.
- Tällöin:
 - (i) Satunnaismuuttujan X ehdollinen jakauma satunnaismuuttujan Y suhteen *ei riipu* ehtomuuttujan Y arvoista.
 - (ii) Satunnaismuuttujan Y ehdollinen jakauma satunnaismuuttujan X suhteen *ei riipu* ehtomuuttujan X arvoista.

Ehdolliset jakaumat ja odotusarvot
Ehdolliset odotusarvot:
Diskreetit jakaumat

- Olkoot satunnaismuuttujat X ja Y *diskreettejä*.
- Satunnaismuuttujan X **ehdollinen odotusarvo** satunnaismuuttujan Y suhteen on satunnaismuuttujan X ehdollisen jakauman odotusarvo:

$$E(X|Y = y) = \sum_x x f_{X|Y}(x|y)$$

- Satunnaismuuttujan Y **ehdollinen odotusarvo** satunnaismuuttujan X suhteen on satunnaismuuttujan Y ehdollisen jakauman odotusarvo:

$$E(Y|X = x) = \sum_y y f_{Y|X}(y|x)$$

Ehdolliset jakaumat ja odotusarvot

Ehdolliset odotusarvot:

Jatkuvat jakaumat

- Olkoot satunnaismuuttujat X ja Y *jatkuvia*.
- Satunnaismuuttujan X **ehdollinen odotusarvo** satunnaismuuttujan Y suhteen on satunnaismuuttujan X ehdollisen jakauman odotusarvo:

$$E(X|Y = y) = \int_{-\infty}^{+\infty} xf_{X|Y}(x|y)dx$$

- Satunnaismuuttujan Y **ehdollinen odotusarvo** satunnaismuuttujan X suhteen on satunnaismuuttujan Y ehdollisen jakauman odotusarvo:

$$E(Y|X = x) = \int_{-\infty}^{+\infty} yf_{Y|X}(y|x)dy$$

Ehdolliset jakaumat ja odotusarvot

Ehdolliset odotusarvot: Kommentteja

- Satunnaismuuttujan X ehdollinen odotusarvo

$$E(X|Y = y)$$

satunnaismuuttujan Y suhteen (ehdolla $Y = y$) *riippuu yleisessä tapauksessa ehtomuuttujan Y arvosta y .*

- Satunnaismuuttujan Y ehdollinen odotusarvo

$$E(Y|X = x)$$

satunnaismuuttujan X suhteen (ehdolla $X = x$) *riippuu yleisessä tapauksessa ehtomuuttujan X arvosta x .*

Riippumattomuus ja ehdolliset odotusarvot 1/2

- Jos satunnaismuuttujat X ja Y ovat *riippumattomia*, ehdolliset odotusarvot *yhtyvät* niiden *reunajakaumien odotusarvoihin*.
- Jos siis X ja Y ovat *riippumattomia*, seuraava pätee:

$$E(X|Y) = E(X)$$

$$E(Y|X) = E(Y)$$

Riippumattomuus ja ehdolliset odotusarvot 2/2

- Olkoot satunnaismuuttujat X ja Y riippumattomia.
- Tällöin:
 - (i) Satunnaismuuttujan X ehdollinen odotusarvo satunnaismuuttujan Y suhteen *ei riipu* ehtomuuttujan Y arvoista.
 - (ii) Satunnaismuuttujan Y ehdollinen odotusarvo satunnaismuuttujan X suhteen *ei riipu* ehtomuuttujan X arvoista.

Ehdolliset odotusarvot satunnaismuuttujina: Iteroidun odotusarvon laki

- Ehdolliset odotusarvot voidaan tulkita *satunnaismuuttujiksi* ehtomuuttujan suhteen.
- Siten satunnaismuuttujan Y ehdollisen odotusarvon odotusarvoksi (satunnaismuuttujan X suhteen) saadaan

$$E_X [E(Y|X)] = E(Y)$$

- Siten satunnaismuuttujan X ehdollisen odotusarvon odotusarvoksi (satunnaismuuttujan Y suhteen) saadaan

$$E_Y [E(X|Y)] = E(X)$$

- Kaavat tunnetaan nimellä **iteroidun odotusarvon laki**.

Iteroidun odotusarvon laki: Perustelu

- Iteroidun odotusarvon laki nähdään oikeaksi *jatkuvan* jakauman tapauksessa seuraavalla tavalla:

$$\begin{aligned} E_X [E(Y|X)] &= \int_{-\infty}^{+\infty} E(Y|X) f_X(x) dx = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} y f_{Y|X}(y|x) dy \right] f_X(x) dx \\ &= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} y f_{Y|X}(y|x) f_X(x) dy \right] dx = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} y f_{XY}(x, y) dy \right] dx \\ &= \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f_{XY}(x, y) dx dy = \int_{-\infty}^{+\infty} y f_Y(y) dy \\ &= E(Y) \end{aligned}$$

Regressiofunktiot ja -käyrät 1/2

- Tarkastellaan satunnaismuuttujan X *ehdollista odotusarvoa*

$$E(X \mid Y = y)$$

ehtomuuttujan Y arvojen *y* *funktiona*.

- Tätä funktiota kutsutaan satunnaismuuttujan X **regressiofunktiksi** satunnaismuuttujan Y suhteen.
- Satunnaismuuttujan X regressiofunktio muuttujan Y suhteen määrittelee **regressiokäyrän**

$$\begin{aligned} x &= g_y(y) \\ &= E(X \mid Y = y) \end{aligned}$$

Regressiofunktiot ja -käyrät 2/2

- Tarkastellaan satunnaismuuttujan Y ehdollista odotusarvoa

$$E(Y | X = x)$$

ehtomuuttujan X arvojen x funktiona.

- Tätä funktiota kutsutaan satunnaismuuttujan Y **regressiofunktiksi** satunnaismuuttujan X suhteen.
- Satunnaismuuttujan Y regressiofunktio muuttujan X suhteen määrittelee **regressiokäyrän**

$$\begin{aligned} y &= g_x(x) \\ &= E(Y | X = x) \end{aligned}$$

Regressiofunktiot ja -käyrät: Kommentteja

- Olkoon

$$x = g_y(y) = E(X | Y = y)$$

satunnaismuuttujan X regressiokäyrä satunnaismuuttujan Y saamien arvojen y suhteen.

- Olkoon

$$y = g_x(x) = E(Y | X = x)$$

satunnaismuuttujan Y regressiokäyrä satunnaismuuttujan X saamien arvojen x suhteen.

- Vaikka funktiot g_y ja g_x ovat funktioina täysin määrättyjä, sattuma määrää, mikä funktioiden arvoista realisoituu.

Regressiokäyrät ja ennustaminen 1/4

- Oletetaan, että satunnaismuuttujien X ja Y yhteisjakauman pistetodennäköisyys- tai tiheysfunktio f_{XY} tunnetaan.
- Haluamme **ennustaa** satunnaismuuttujan X (tai Y) arvon satunnaismuuttujan Y (tai X) saaman arvon perusteella.

Regressiokäyrät ja ennustaminen 2/4

- Tehtävä 1:
 - (i) Haluamme *ennustaa* satunnaismuuttujan X arvon satunnaismuuttujan Y saaman arvon perusteella.
 - (ii) Olkoon ennustettu arvo $d(X | Y)$.
 - (iii) Miten ennuste $d(X | Y)$ valitaan *optimaalisella tavalla*?

- Tehtävä 2:
 - (i) Haluamme *ennustaa* satunnaismuuttujan Y arvon satunnaismuuttujan X saaman arvon perusteella.
 - (ii) Olkoon ennustettu arvo $d(Y | X)$.
 - (iii) Miten ennuste $d(Y | X)$ valitaan *optimaalisella tavalla*?

Regressiokäyrät ja ennustaminen 3/4

- Tehtävän 1 ratkaisu:

Valitaan $d(X | Y)$ siten, että *ennusteen keskineliövirhe*

$$E[X - d(X | Y)]^2$$

minimoituu.

- Voidaan osoittaa, että keskineliövirhe minimoituu valinnalla

$$d(X | Y) = E(X | Y)$$

- Siten *ehdollinen odotusarvo* $E(X | Y)$ on *keskineliövirheen mielessä optimaalinen ennuste* satunnaismuuttujan X saamille arvoille.

Regressiokäyrät ja ennustaminen 4/4

- Tehtävän 2 ratkaisu:

Valitaan $d(Y | X)$ siten, että *ennusteen keskineliövirhe*

$$E[Y - d(Y | X)]^2$$

minimoituu.

- Voidaan osoittaa, että keskineliövirhe minimoituu valinnalla

$$d(Y | X) = E(Y | X)$$

- Siten *ehdollinen odotusarvo* $E(Y | X)$ on *keskineliövirheen mielessä optimaalinen ennuste* satunnaismuuttujan Y saamille arvoille.

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat:

1. esimerkki nopanheitosta 1/6

- Heitetään *virheetöntä* noppaa kaksi kertaa.
- Määritellään *satunnaismuuttujat* X ja Y :
 $X = \text{tulos (silmäluku) 1. heitosta}$
 $Y = \text{tulos (silmäluku) 2. heitosta}$
- Satunnaismuuttujien X ja Y mahdolliset arvot:
 $X: \{1, 2, 3, 4, 5, 6\}$
 $Y: \{1, 2, 3, 4, 5, 6\}$
- Määrätään satunnaismuuttujien X ja Y *ehdolliset jakaumat*.

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 1. esimerkki nopanheitosta 2/6

- Satunnaismuuttujien X ja Y yhteisjakauma ja reunajakaumat:

Todennäköisyydet $\Pr(X = x, Y = y)$

	1	2	3	4	5	6	Yht
2. heiton silmäluku y							
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
Yht	1/6	1/6	1/6	1/6	1/6	1/6	1

1. heiton silmäluku x

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 1. esimerkki nopanheitosta 3/6

- Kalvon 2/6 satunnaismuuttujan X ehdolliset jakaumat saadaan jakamalla yhteisjakauman todennäköisyydet antavassa taulukossa kunkin rivin todennäköisyydet vastaavilla rivisummilla eli vastaavilla reunatodennäköisyyksillä.
- Esimerkki:

Satunnaismuuttujan X ehdollinen jakauma satunnaismuuttujan Y suhteen, kun $Y = 3$:

$$f_{X|Y}(x|Y=3) = \frac{f_{XY}(x,3)}{f_Y(3)} = \frac{1/36}{1/6} = \frac{1}{6}; x = 1, 2, 3, 4, 5, 6$$

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 1. esimerkki nopanheitosta 4/6

- Satunnaismuuttujan X ehdolliset jakaumat ovat taulukon riveinä:

Todennäköisyydet $\Pr(X = x \mid Y = y)$

	1	2	3	4	5	6	Yht
2. heiton silmäluku y							
6	1/6	1/6	1/6	1/6	1/6	1/6	1
5	1/6	1/6	1/6	1/6	1/6	1/6	1
4	1/6	1/6	1/6	1/6	1/6	1/6	1
3	1/6	1/6	1/6	1/6	1/6	1/6	1
2	1/6	1/6	1/6	1/6	1/6	1/6	1
1	1/6	1/6	1/6	1/6	1/6	1/6	1
Yht	1	1	1	1	1	1	

1. heiton silmäluku x

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 1. esimerkki nopanheitosta 5/6

- Kalvon 2/6 satunnaismuuttujan Y ehdolliset jakaumat saadaan jakamalla yhteisjakauman todennäköisyydet antavassa taulukossa kunkin sarakkeen todennäköisyydet vastaavilla sarakesummilla eli vastaavilla reunatodennäköisyyksillä.
- Esimerkki:

Satunnaismuuttujan Y ehdollinen jakauma satunnaismuuttujan X suhteen, kun $X = 4$:

$$f_{Y|X}(y|X = 4) = \frac{f_{XY}(4, y)}{f_X(4)} = \frac{1/36}{1/6} = \frac{1}{6}; y = 1, 2, 3, 4, 5, 6$$

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 1. esimerkki nopanheitosta 6/6

- Satunnaismuuttujan Y ehdolliset jakaumat ovat taulukon sarakkeina:

Todennäköisyydet $\Pr(Y = y | X = x)$

	1	2	3	4	5	6	Yht
2. heiton silmäluku y							
6	1/6	1/6	1/6	1/6	1/6	1/6	1
5	1/6	1/6	1/6	1/6	1/6	1/6	1
4	1/6	1/6	1/6	1/6	1/6	1/6	1
3	1/6	1/6	1/6	1/6	1/6	1/6	1
2	1/6	1/6	1/6	1/6	1/6	1/6	1
1	1/6	1/6	1/6	1/6	1/6	1/6	1
Yht	1	1	1	1	1	1	

1. heiton silmäluku x

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 2. esimerkki nopanheitosta 1/4

- Heitetään *virheetöntä* noppaa kaksi kertaa.
- Määritellään *satunnaismuuttujat* X , Y ja Z :
 - $X = \text{tulos (silmäluku) 1. heitosta}$
 - $Y = \text{tulos (silmäluku) 2. heitosta}$
 - $Z = X + Y = \text{silmälukujen summa}$
- Satunnaismuuttujien X , Y ja Z mahdolliset arvot:
 - $X: \{1, 2, 3, 4, 5, 6\}$
 - $Y: \{1, 2, 3, 4, 5, 6\}$
 - $Z: \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
- Määrätään satunnaismuuttujan Z *ehdolliset jakaumat* satunnaismuuttujan X suhteen.

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 2. esimerkki nopanheitosta 2/4

- Satunnaismuuttujien X ja $Z = X + Y$ yhteisjakauma ja reunajakaumat:

1. nopan silmäluku x

	1	2	3	4	5	6	Yht
Silmälukujen summa z	12	0	0	0	0	1/36	1/36
	11	0	0	0	1/36	1/36	2/36
	10	0	0	1/36	1/36	1/36	3/36
	9	0	1/36	1/36	1/36	1/36	4/36
	8	1/36	1/36	1/36	1/36	1/36	5/36
	7	1/36	1/36	1/36	1/36	1/36	6/36
	6	1/36	1/36	1/36	1/36	0	5/36
	5	1/36	1/36	1/36	1/36	0	4/36
	4	1/36	1/36	1/36	0	0	3/36
	3	1/36	1/36	0	0	0	2/36
	2	1/36	0	0	0	0	1/36
	Yht	1/6	1/6	1/6	1/6	1/6	1/6

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 2. esimerkki nopanheitosta 3/4

- Kalvon 2/4 satunnaismuuttujan Z ehdolliset jakaumat saadaan jakamalla yhteisjakauman todennäköisyydet antavassa taulukossa kunkin sarakkeen todennäköisyydet vastaavilla sarakesummilla eli vastaavilla reunatodennäköisyyksillä.
- Esimerkkejä:
 - (i) Tapahtuman $Z = 8$ ehdollinen todennäköisyys, kun ehtotapahtumana on $X = 3$:

$$\Pr(Z = 8|X = 3) = f_{Z|X}(8|X = 3) = \frac{f_{ZX}(8,3)}{f_X(3)} = \frac{1/36}{1/6} = \frac{1}{6}$$

- (ii) Tapahtuman $Z = 10$ ehdollinen todennäköisyys, kun ehtotapahtumana on $X = 3$:

$$\Pr(Z = 10|X = 3) = f_{Z|X}(10|X = 3) = \frac{f_{ZX}(10,3)}{f_X(3)} = \frac{0}{1/6} = 0$$

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset jakaumat: 2. esimerkki nopanheitosta 4/4

- Satunnaismuuttujan Z ehdolliset jakaumat ovat taulukon sarakkeina:

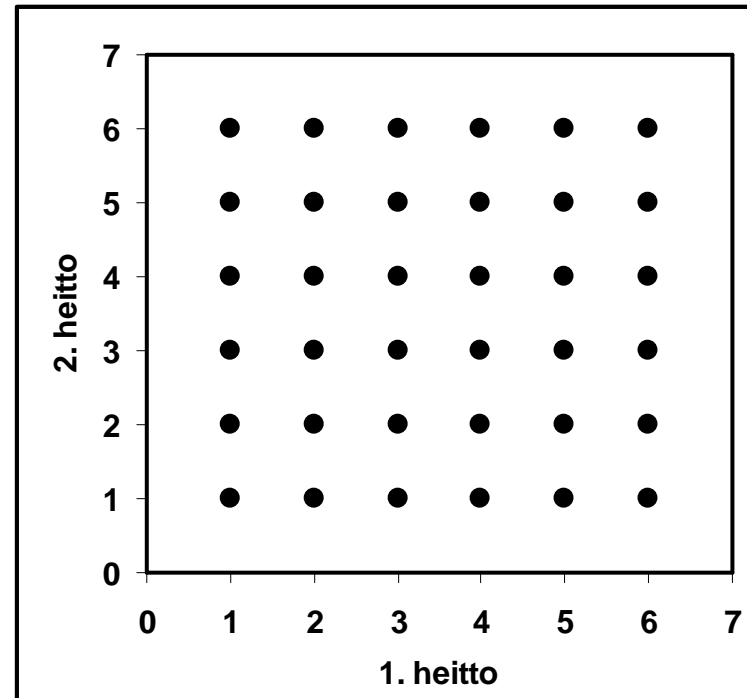
1. nopan silmäluku x

	1	2	3	4	5	6	Yht	
Silmälukujen summa z	12	0	0	0	0	0	1/6	1/6
	11	0	0	0	0	1/6	1/6	2/6
	10	0	0	0	1/6	1/6	1/6	3/6
	9	0	0	1/6	1/6	1/6	1/6	4/6
	8	0	1/6	1/6	1/6	1/6	1/6	5/6
	7	1/6	1/6	1/6	1/6	1/6	1/6	6/6
	6	1/6	1/6	1/6	1/6	1/6	0	5/6
	5	1/6	1/6	1/6	1/6	0	0	4/6
	4	1/6	1/6	1/6	0	0	0	3/6
	3	1/6	1/6	0	0	0	0	2/6
	2	1/6	0	0	0	0	0	1/6
	Yht	1	1	1	1	1	1	

Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset odotusarvot: 1. esimerkki nopanheitosta 1/2

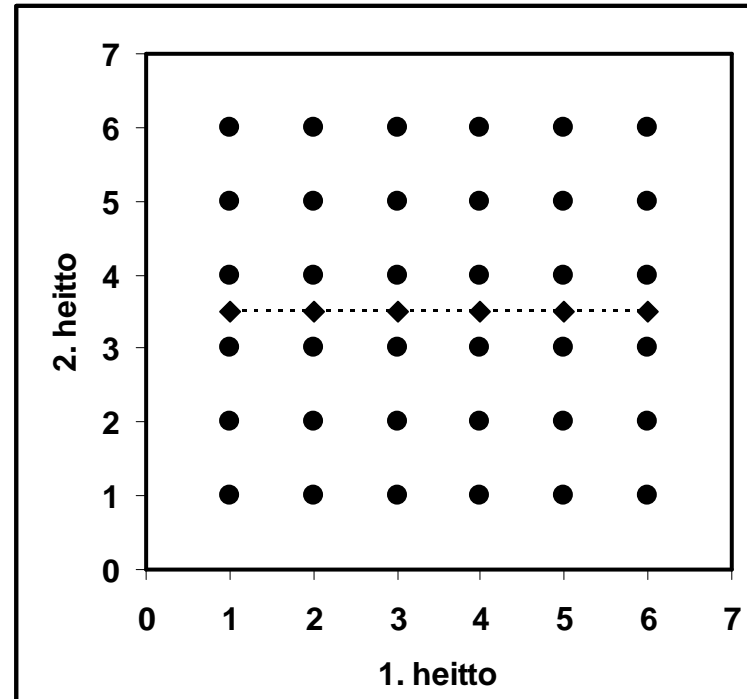
- Heitetään *virheetöntä* noppaa kaksi kertaa:
 $X =$ tulos 1. heitosta
 $Y =$ tulos 2. heitosta
- Kuva oikealla havainnollistaa satunnaismuuttujien X ja Y yhteisjakaumaa.
- Kuvassa on 36 pistettä, joista jokaisen todennäköisyys on $1/36$.



Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset odotusarvot: 1. esimerkki nopanheitosta 2/2

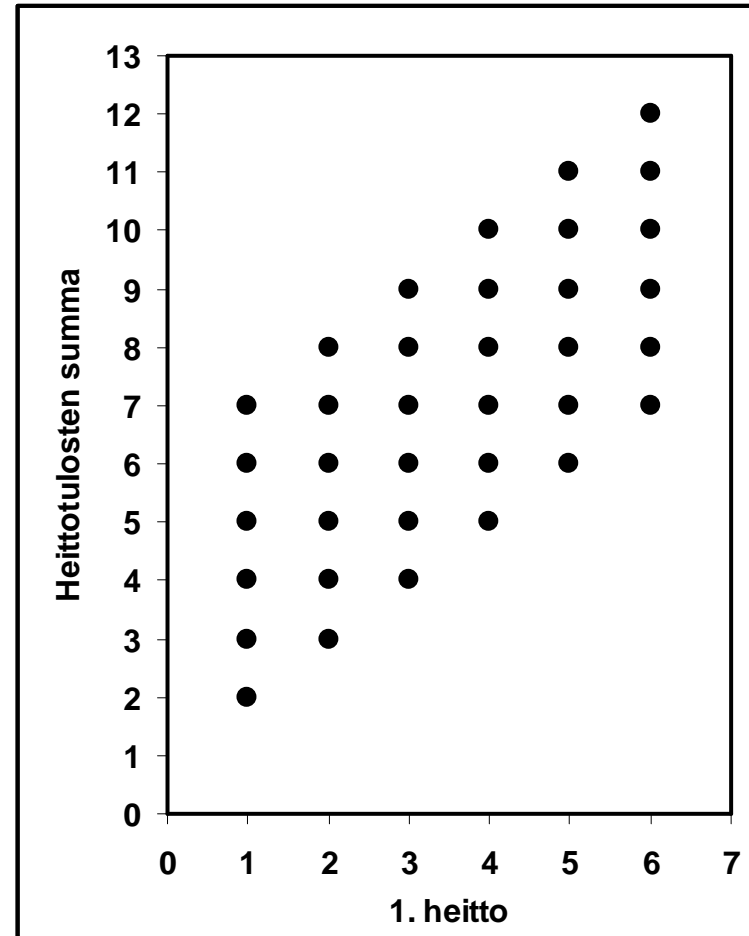
- Satunnaismuuttujan X (= tulos 1. heitosta) ehdollisia odotusarvoja satunnaismuuttujan Y (= tulos 2. heitosta) arvojen suhteen on merkitty katkoviivan yhdistämällä vinoneliöillä.
- 1. nopanheiton tuloksen tuntemisesta *ei ole hyötyä* 2. nopanheiton tuloksen ennustamisessa, koska heittojen tulokset ovat riippumattomia.



Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset odotusarvot: 2. esimerkki nopanheitosta 1/2

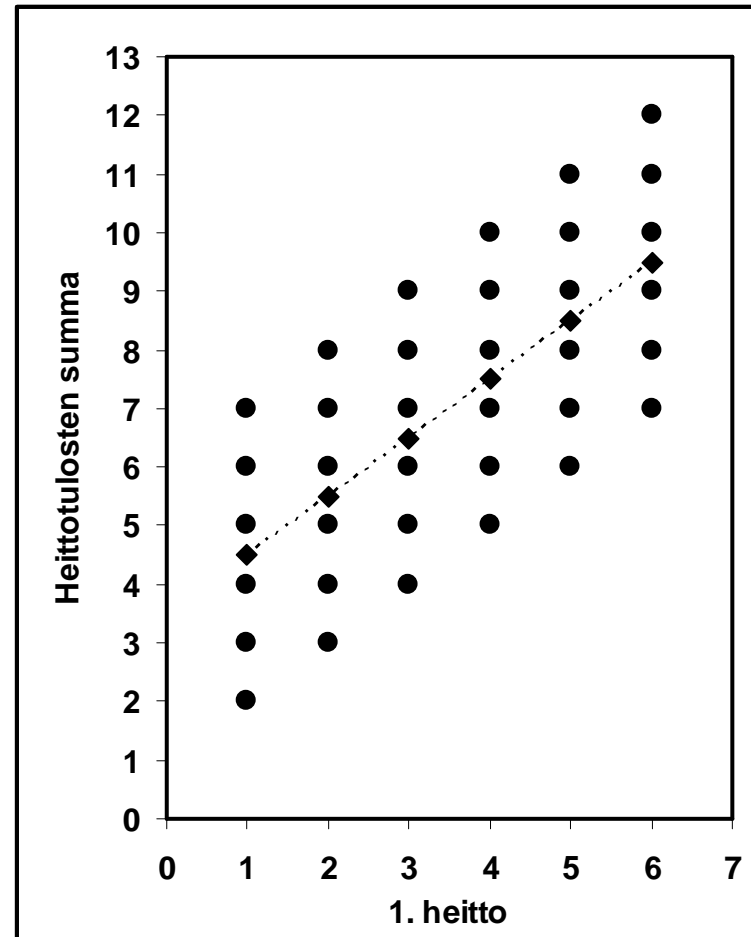
- Heitetään *virheetöntä* noppaa kaksi kertaa:
 $X =$ tulos 1. heitosta
 $Y =$ tulos 2. heitosta
 $Z = X + Y$
- Kuva oikealla havainnollistaa satunnaismuuttujien X ja Z yhteisjakaumaa.
- Kuvassa on 36 pistettä, joista jokaisen todennäköisyys on $1/36$.



Ehdolliset jakaumat ja odotusarvot

Diskreetin 2-ulotteisen jakauman ehdolliset odotusarvot: 2. esimerkki nopanheitosta 2/2

- Satunnaismuuttujan Z (= heittotulosten summa) *ehdollisia odotusarvoja* satunnaismuuttujan X (= tulos 1. heitosta) arvojen suhteen on merkitty katkoviivan yhdistämällä *vinoneliöillä*.
- 1. nopanheiton tuloksen tuntemisesta *on hyötyä* heittotulosten summan *ennustamisessa*, koska summa *ei ole riippumaton* 1. heiton tuloksesta.



Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkki 1/2

- Valitaan luvut X ja Y *satunnaisesti* väliltä $[0, 1]$ kahdessa vaiheessa:

- (1) Valitaan väliltä $(0, 1)$ *satunnaisesti* luku X :

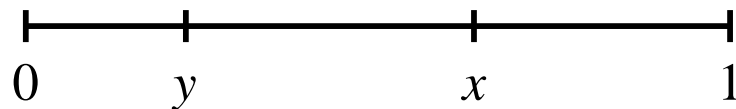
$$X \sim \text{Uniform}(0, 1)$$

Olkoon valittu luku x .

- (2) Valitaan väliltä $(0, x)$ *satunnaisesti* luku Y :

$$Y \sim \text{Uniform}(0, x)$$

Olkoon valittu luku y .



- Tehtävänä on määrätä satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktio $f_{XY}(x, y)$.

Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkki 2/2

- Oletuksien mukaan satunnaismuuttujan X (*reuna-*) *jakauma* tunnetaan:

$$f_X(x) = 1 \quad 0 < x < 1$$

$$f_X(x) = 0 \quad \text{muulloin}$$

- Oletuksien mukaan satunnaismuuttujan Y *ehdollinen jakauma* ehdolla $X = x$ tunnetaan:

$$f_{Y|X}(y | x) = 1/x \quad 0 < y < x$$

$$f_{Y|X}(y | x) = 0 \quad \text{muulloin}$$

Ehdolliset jakaumat ja odotusarvot

Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkin yhteisjakauma ja Y :n reunajakauma

- Satunnaismuuttujien X ja Y yhteisjakauma on

$$\begin{aligned} f_{XY}(x, y) &= f_{Y|X}(y | x)f_X(x) \\ &= 1/x \qquad 0 < y < x < 1 \end{aligned}$$

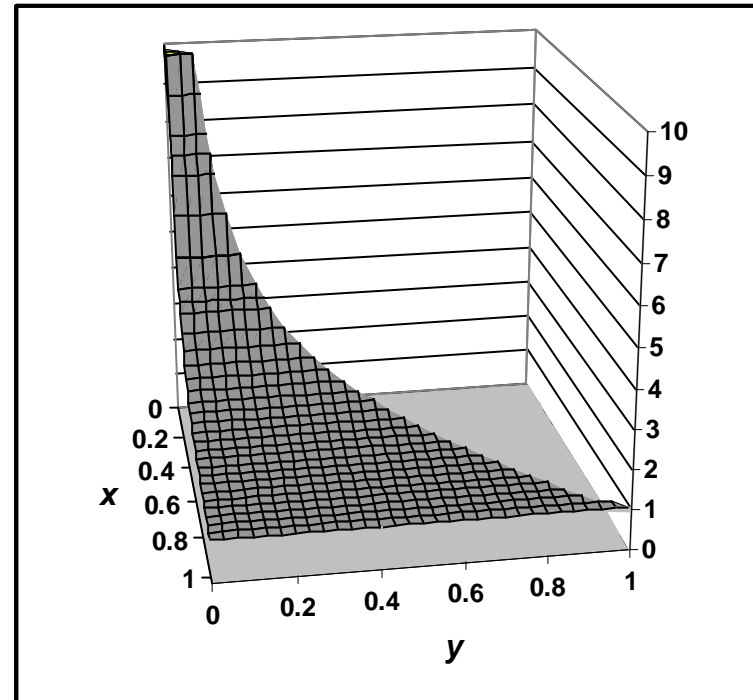
- Satunnaismuuttujan Y reunajakauma on

$$f_Y(y) = \int_y^1 f_{XY}(x, y)dx = \int_y^1 \frac{1}{x} dx = [\log x]_y^1 = -\log y$$

Ehdolliset jakaumat ja odotusarvot

Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkin yhteisjakauman tiheysfunktio 1/2

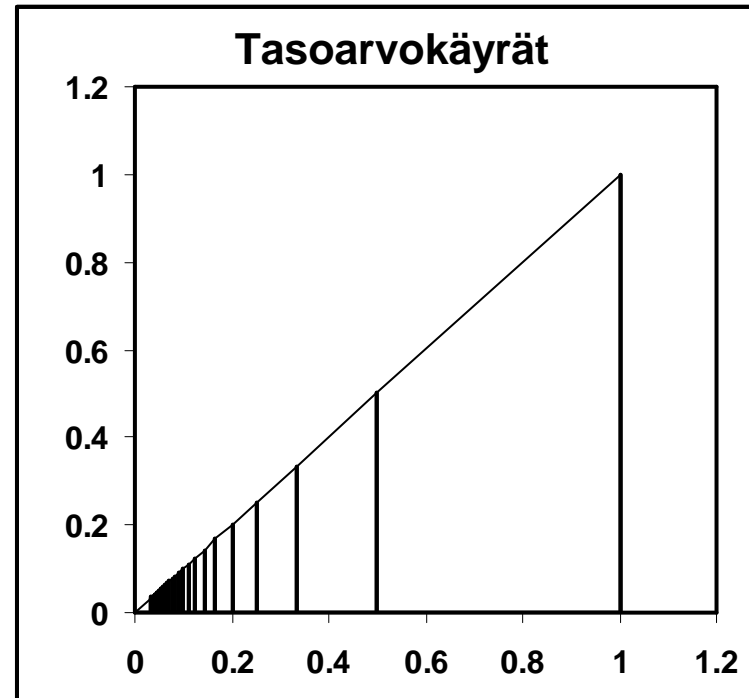
- $X \sim \text{Uniform}(0, 1)$
- $Y | X = x \sim \text{Uniform}(0, x)$
- Kuva oikealla esittää satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktioita
 $f_{XY}(x, y) = 1/x, 0 < y < x < 1$



Ehdolliset jakaumat ja odotusarvot

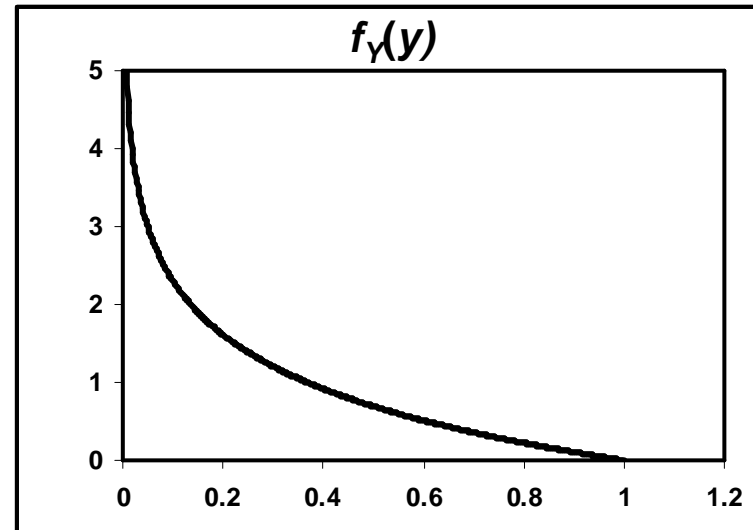
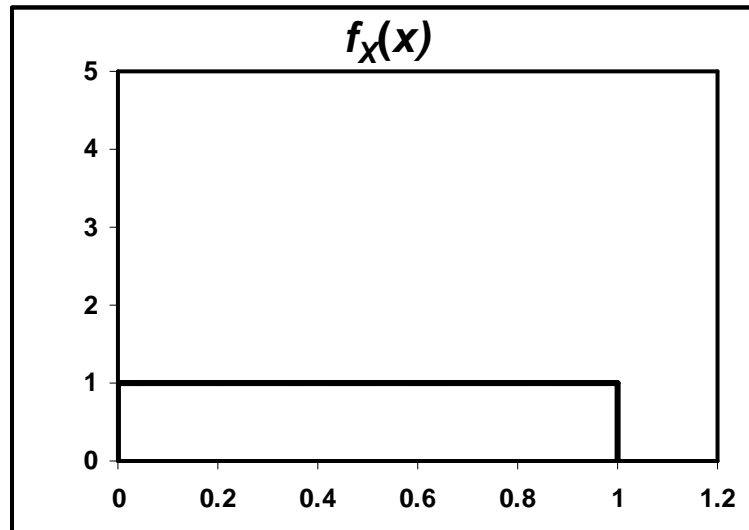
Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkin yhteisjakauman tiheysfunktio 2/2

- $X \sim \text{Uniform}(0, 1)$
- $Y | X = x \sim \text{Uniform}(0, x)$
- Kuva oikealla esittää satunnaismuuttujien X ja Y yhteisjakauman tiheysfunktion $f_{XY}(x, y) = 1/x$, $0 < y < x < 1$ tasa-arvokäyriä $x = 1/k$, $0 < y < x$ kun $k = 1, 2, \dots, 30$.
- Kuvaan on merkitty myös suora $y = x$, $0 < x < 1$



Ehdolliset jakaumat ja odotusarvot

Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkin reunajakaumien tiheysfunktiot



- Kuvat yllä esittävät satunnaismuuttujien X ja Y *reunajakaumien tiheysfunktioita*:

$$f_X(x) = 1, 0 < x < 1$$

$$f_Y(y) = -\log(y), 0 < y < x < 1$$

Ehdolliset jakaumat ja odotusarvot

Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkin Y :n ehdollinen odotusarvo $1/2$

- Oletetaan, että

$$X \sim \text{Uniform}(0, 1)$$

$$Y|X = x \sim \text{Uniform}(0, x)$$

- Koska X ja $Y|X = x$ noudattavat tasaista jakaumaa, niin

$$E(X) = \frac{1}{2}$$

$$E(Y|X = x) = \frac{1}{2}x$$

- Satunnaismuuttujan Y ehdollisen odotusarvon $E(Y|X)$ odotusarvo satunnaismuuttujan X suhteen on iteroidun odotusarvon lain mukaan

$$E(Y) = E_x [E(Y|X)] = \frac{1}{2}E(X) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Ehdolliset jakaumat ja odotusarvot

Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkin Y :n ehdollinen odotusarvo 2/2

- Koska toisaalta

$$E(Y) = \int_0^1 y f_Y(y) dy$$

ja

$$f_Y(y) = -\log(y)$$

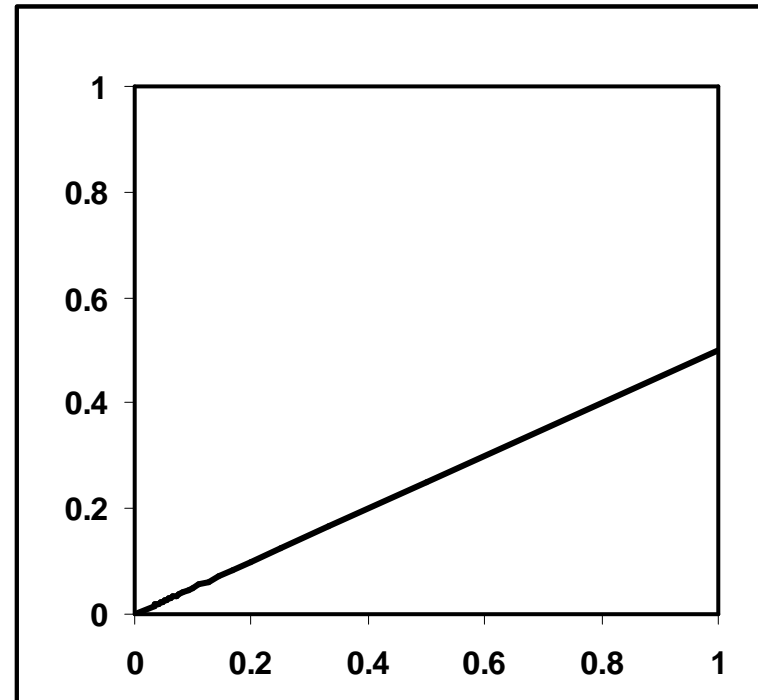
olemme todistaneet seuraavan integraalikaavan:

$$-\int_0^1 y \log(y) dy = \frac{1}{4}$$

Ehdolliset jakaumat ja odotusarvot

Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkin Y :n regressiokäyrä X :n suhteen $1/2$

- $X \sim \text{Uniform}(0, 1)$
- $Y | X = x \sim \text{Uniform}(0, x)$
- Satunnaismuuttujan Y ehdollinen odotusarvo satunnaismuuttujan X suhteen on
$$E(Y|X = x) = x/2$$
- Siten muuttujan y regressiokäyrä muuttujan x suhteen on muotoa
$$y = x/2$$



Ehdolliset jakaumat ja odotusarvot

Jatkuvan 2-ulotteisen jakauman ehdolliset jakaumat: Esimerkin Y :n regressiokäyrä X :n suhteen 2/2

- Huomaa, että satunnaismuuttujan X saaman arvon tuntemisesta *on* tässä tapauksessa *hyötyä* satunnaismuuttujan Y saaman arvon *ennustamisessa*.

