Johdatus todennäköisyyslaskentaan
Kevät 2014

Luento 11 / 13

Jukka Kohonen
Matematiikan ja tilastotieteen laitos
Helsingin yliopisto

VARIANSSI

Jakauman leveys

- Usein halutaan tunnusluku, joka kertoo miten tiiviisti mahdolliset arvot (eli X:n jakauma) **keskittyvät** odotusarvon lähelle
  - esim. ”koko arvojoukon leveys” ei oikein kerro tätä asiaa (vrt. tasa- ja kolmiojakauma)
- Eräs ratkaisu: **varianssi** (Tuominen s. 82)

  $$\mathbf{Var(X) = E((X - \mu)^2)}$$

  ja varianssin neliöjuuri eli **hajonta**
  - varianssilla paljon käteviä **laskusääntöjä**
  - varianssi myös **tunnetaan** useille tutuille jakaumille
- Varianssille käytetään vaihtelevasti merkintöjä Var(X) ja $D^2(X)$, ne tarkoittavat samaa.

Varianssin idea

- Oletetaan tunnetuksi $E(X) = \mu$.
- Tiedetään, ettei X aina (eikä edes kovin usein) osu odotusarvoonsa. (Ehkä ei koskaan)
- Katsotaan, paljonko se meni huti (= **poikkeama**) ja korotetaan poikkeama toiseen (= **neliöpoikkeama**)
- Jokaiseen X:n mahdolliseen arvoon liittyy vastaava neliöpoikkeama $(X-\mu)^2$
- X:n mahdollisilla arvoilla on todennäköisyydet, joten voidaan laskea, miten neliöpoikkeama on jakautunut. Lasketaan sen odotusarvo ja nimitetään sitä varianssiksi.
- Suuri varianssi siis merkitsee, että X usein poikkeaa paljon odotusarvostaan
- Pienimillään varianssi voisi olla 0, jos X on aina $= \mu$

Varianssin laskentaa

- Lasketaan suoraan määritelmän perusteella
  - Reilun kolikon varianssi
  - Epäreilun kolikon varianssi

- Sitten huomataan kirjasta (s. 82) lause 3.2.3
  - Miksi se pitää paikkansa?
  - Miten sitä käytetään?

- Lauseen 3.2.3 avulla laskenta usein helpompaa, koska lausekkeet sievempiä

Reilun kolikon varianssi

Diskreetti satunnaismuuttuja joukossa {0, 1},

$P(X=0) = ½$

$P(X=1) = ½$

Odotusarvo: $\mu = E(X) = ½ \cdot 0 + ½ \cdot 1 =$ **½**

Poikkeama	$(X-\mu)$	on joko $+½$ tai $-½$
Neliöpoikkeama	$Z = (X-\mu)^2$	on varmasti ¼
$Var(X) =$	$E(Z)$	$= ¼$

Epäreilun kolikon varianssi

Diskreetti satunnaismuuttuja joukossa {0, 1},

$P(X=0) = q$

$P(X=1) = p \qquad p+q=1$

jos X=0 jos X=1

Odotusarvo: $\mu = E(X) = q \cdot 0 + p \cdot 1 = \mathbf{p}$

Poikkeama $(X-\mu)$ on joko $-p$ tai $(1-p)$

Neliöpoikkeama $Z = (X-\mu)^2$ on joko $(-p)^2$ tai $(1-p)^2$

$\mathrm{Var}(X) = E(Z) = q \cdot (-p)^2 + p \cdot (1-p)^2$

$= qp^2 + pq^2$

$= \mathbf{pq}.$

Nopan varianssi

- (Lasketaan taululla jos jaksetaan)

Lause 3.2.3

- Merkitään $\mu = E(X)$
- Varianssin määritelmästä lähdetään laskemaan binomin neliötä auki...

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

Nopanheitto lauseella 3.2.3

Arvojoukko {1,2,3,4,5,6}, symmetrisesti

$$
\begin{aligned}
E(X) &= 3.5 \\
Var(X) &= E(X^2) - [E(X)]^2 \\
&= (1+4+9+16+25+36) / 6 - 3.5^2 \\
&= (1+4+9+16+25+36) / 6 - 3.5^2 \\
&= 91/6 - 3.5^2 \\
&\approx 2.92
\end{aligned}
$$

Tasajakauman varianssi

- $X \sim$ Tas(0, 1)
- Integroidaan taululla, tuloksena pitäisi olla 1/12

MONEN MUUTTUJAN SUMMA JA NORMAALIJAKAUMA

Esimerkki: **Neljän** tasajakautuneen summa

(Herra K matkustaa kahdella bussilla peräkkäin)

odotus1	$\boldsymbol{X}$	$\sim \text{Tas}(0, 4)$	$E(X) = 2$
ajoaika1	$\boldsymbol{Y}$	$\sim \text{Tas}(10, 14)$	$E(Y) = 12$
odotus2	$\boldsymbol{Z}$	$\sim \text{Tas}(0, 6)$	$E(Z) = 3$
ajoaika2	$\boldsymbol{W}$	$\sim \text{Tas}(20, 26)$	$E(W) = 23$

Matka-aika $\boldsymbol{M} = X+Y+Z+W$ $E(M) = 40$

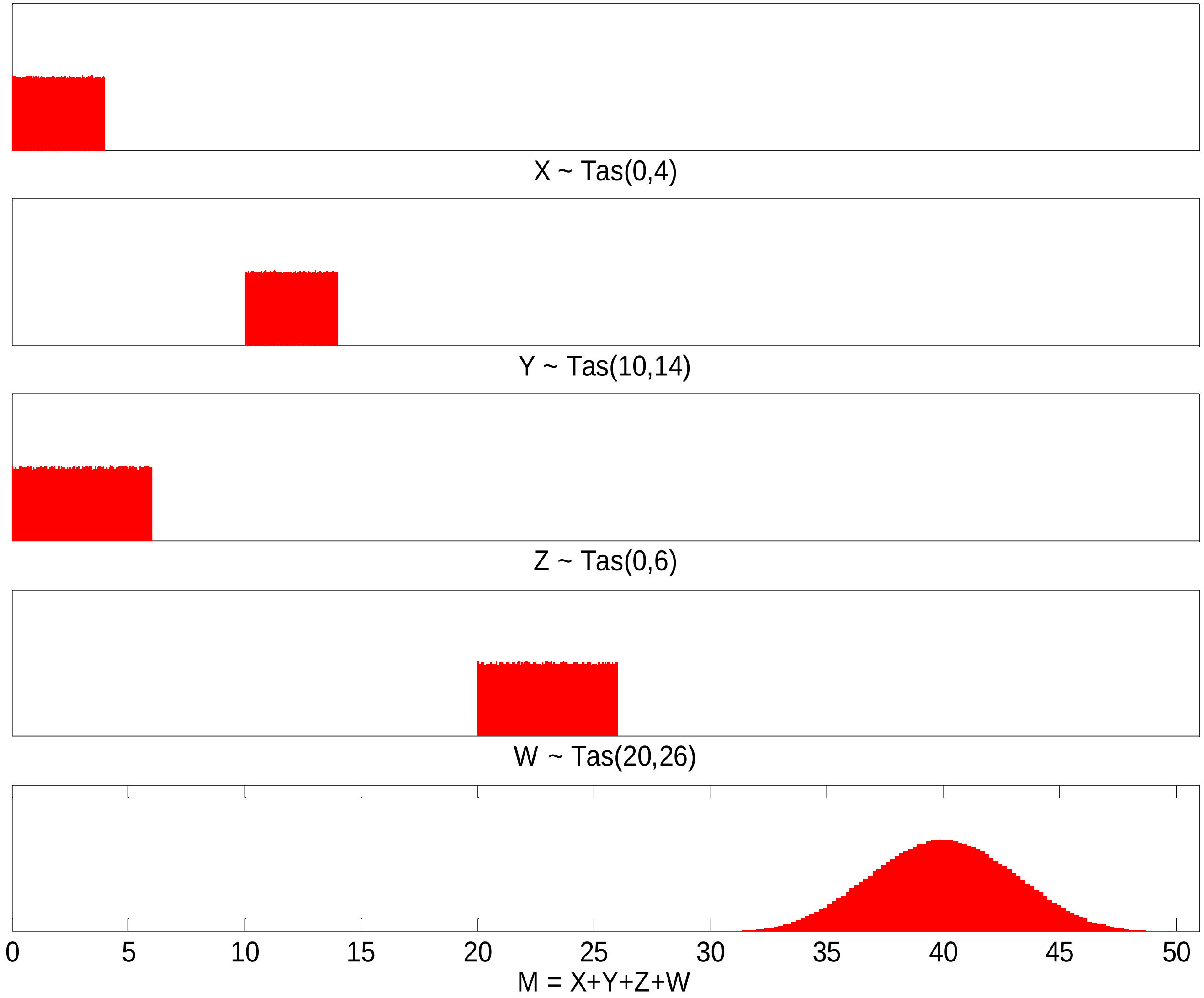
- *M* vähintään $0+10+0+20 = 30$
- *M* enintään $4+14+6+26 = 50$

Minkä muotoinen jakauma *M*:llä on?

Ei varmaan tasajakauma, vaikka onkin välillä (30, 50) ja vaikka odotusarvo 40 on välin keskellä.

Esimerkki: **Neljän** tasajakautuneen summa

(Herra K matkustaa matkustaa kahdella bussilla)



Tutun näköinen jakauma?