

SF1901 Sannolikhetsteori och statistik I

Jimmy Olsson

Föreläsning 6
15 november 2016



Idag

Förra gången

Mer om väntevärden och varianser (Kap. 5.2–5.3)

Beroendemått (Kap. 5.4)

Summor, linjärkombinationer och stora talens lag (Kap. 5.5–5.6)



Idag

Förra gången

Mer om väntevärden och varianser (Kap. 5.2–5.3)

Beroendemått (Kap. 5.4)

Summor, linjärkombinationer och stora talens lag (Kap. 5.5–5.6)



Varians och standardavvikelse för stokastiska variabler

- ▶ *Variansen* för en stokastisk variabel X definieras som

$$\mathbb{V}(X) = \mathbb{E}[(X - \mu_X)^2],$$

där $\mu_X = \mathbb{E}(X)$, dvs.

$$\mathbb{V}(X) = \begin{cases} \sum (k - \mu_X)^2 p_X(k) & \text{om } X \text{ är diskret,} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx & \text{om } X \text{ är kontinuerlig.} \end{cases}$$

- ▶ Variansen mäter hur mycket X avviker från $\mathbb{E}(X)$ i medeltal och ger sålunda ett mått på fördelningens *spridning*.
- ▶ Ibland betecknas variansen även som σ_X^2 .
- ▶ Man visar enkelt att $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.
- ▶ Exempel!



Flerdimensionella s.v.

- ▶ En n -dimensionell s.v. (X_1, \dots, X_n) är en \mathbb{R}^n -värd funktion definierad på ett utfallsrum Ω .
- ▶ Den simultana fördelningsfunktionen definieras i detta fall som

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n),$$
$$(x_1, \dots, x_n) \in \mathbb{R}^n.$$



Flerdimensionella sannolikhets- och täthetsfunktioner

- ▶ Funktionen

$$p_{X,Y}(j, k) = \mathbb{P}(X = j, Y = k), \quad j, k \in \{0, 1, 2, \dots\}$$

kallas *sannolikhetsfunktionen* för den diskreta tvådimensionella s.v. (X, Y) .

- ▶ Om det finns en icke-negativ funktion $f_{X,Y}(x, y)$ sådan att

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

för alla A sägs den tvådimensionella s.v. (X, Y) vara *kontinuerlig*. Funktionen $f_{X,Y}(x, y)$ kallas *täthetsfunktionen* för (X, Y) .

- ▶ Motsvarande definitioner gäller i det flerdimensionella fallet.



Oberoende s.v.

Definition

De s.v. X och Y kallas oberoende om

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

för alla mängder A och B .

- ▶ Det visade sig att

$$\begin{aligned} X \text{ och } Y \text{ oberoende} &\Leftrightarrow F_{X,Y}(x,y) = F_X(x)F_Y(y) \\ \Leftrightarrow \begin{cases} p_{X,Y}(j,k) = p_X(j)p_Y(k) & \text{för alla } j, k \quad (\text{disk. fallet}), \\ f_{X,Y}(x,y) = f_X(x)f_Y(y) & \text{för alla } x, y \quad (\text{kont. fallet}). \end{cases} \end{aligned}$$

- ▶ Det visade sig tämligen enkelt att bestämma fördelningsfunktionerna för $\max\{X_1, X_2, \dots, X_n\}$ och $\min\{X_1, X_2, \dots, X_n\}$ då X_j :na var oberoende och likafördelade.



Idag

Förra gången

Mer om väntevärden och varianser (Kap. 5.2–5.3)

Beroendemått (Kap. 5.4)

Summor, linjärkombinationer och stora talens lag (Kap. 5.5–5.6)



Varians för affin transformation av en s.v.

- ▶ Exempel!
- ▶ Med hjälp av definitionen av varians visar man enkelt följande.

Sats

För alla s.v. X och konstanter a och b gäller att

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b,$$

$$\mathbb{V}(aX + b) = a^2\mathbb{V}(X),$$

$$\mathbb{D}(aX + b) = |a|\mathbb{D}(X).$$

- ▶ OBS! Konstanten b bidrar inte till variansen.



Väntevärde för en funktion av flerdimensionell s.v.

- ▶ Låt (X, Y) vara en tvådimensionell s.v. och $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ en reellvärd funktion. Sätt $Z = g(X, Y)$. Då gäller följande.

Sats

$$\mathbb{E}(Z) = \begin{cases} \sum_{j,k} g(j, k) p_{X,Y}(j, k) & \text{om } (X, Y) \text{ är diskret,} \\ \iint_{\mathbb{R}^2} g(x, y) f_{X,Y}(x, y) dx dy & \text{om } (X, Y) \text{ är kontinuerlig.} \end{cases}$$

- ▶ Motsvarande gäller i det flerdimensionella fallet.
- ▶ Som vi ska se i det följande kan ovan sats användas för att erhålla ett generellt resultat för det viktiga fallet då X och Y är oberoende och

$$g(X, Y) = XY.$$



Väntevärde för produkt av oberoende s.v.

Sats

Om X och Y är oberoende gäller att

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

- ▶ OBS! Resultatet gäller i allmänhet *inte* om X och Y beroende.
- ▶ Ovan resultat kan generaliseras till fler än två s.v.

Sats

Om X_1, X_2, \dots, X_n är oberoende gäller att

$$\mathbb{E}(X_1 X_2 \cdots X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2) \cdots \mathbb{E}(X_n).$$



Väntevärde för linjärkombination av s.v.

- ▶ Låt X_1, \dots, X_n vara s.v. och betrakta en linjärkombination

$$a_1X_1 + a_2X_2 + \dots + a_nX_n + b = \sum_{i=1}^n a_iX_i + b,$$

där konstanterna a_1, \dots, a_n och b kan vara positiva eller negativa tal.

- ▶ Följande sats visas på samma sätt som den föregående och gäller även om X_i :na är beroende.

Sats

För alla s.v. X_1, \dots, X_n och konstanter a_1, \dots, a_n , b gäller att

$$\mathbb{E} \left(\sum_{i=1}^n a_iX_i + b \right) = \sum_{i=1}^n a_i\mathbb{E}(X_i) + b.$$



Idag

Förra gången

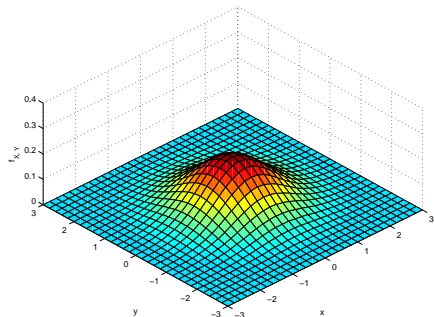
Mer om väntevärden och varianser (Kap. 5.2–5.3)

Beroendemått (Kap. 5.4)

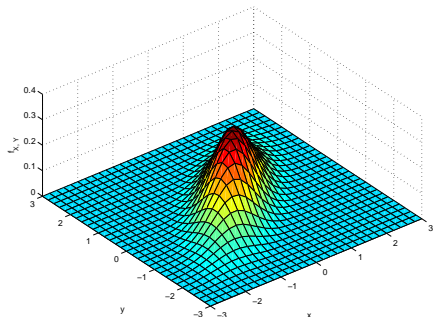
Summor, linjärkombinationer och stora talens lag (Kap. 5.5–5.6)



Exempel: täthetsfunktion för oberoende/beroende s.v.



(a) Oberoende variabler



(b) Beroende variabler

Beroendemått

- ▶ Ett *beroendemått* sammanfattar, i ett enda tal, hur två s.v. X och Y samvarierar.
- ▶ Vi kommer att diskutera två vanliga och närbesläktade beroendemått, nämligen *kovarians* och *korrelation*.



Kovarians

- ▶ Kovariansen mellan två s.v. X och Y definieras som

$$\mathbb{C}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)].$$

- ▶ Notera att $\mathbb{C}(X, X) = \mathbb{E}[(X - \mu_X)^2] = \mathbb{V}(X)$.
- ▶ Vidare ger symmetri att $\mathbb{C}(X, Y) = \mathbb{C}(Y, X)$.
- ▶ Om någon av variablerna är konstant blir $\mathbb{C}(X, Y) = 0$.
- ▶ Intuition: kovariansen bör bli
 - ▶ *positiv* om det finns en tendens att $(X - \mu_X)(Y - \mu_Y)$ är positiv, dvs. att variablerna avviker *åt samma håll* från sina respektive väntevärden,
 - ▶ *negativ* om det finns en tendens att $(X - \mu_X)(Y - \mu_Y)$ är negativ, dvs. att variablerna avviker *åt olika håll* från sina respektive väntevärden.



Kovarians (forts.)

- ▶ Följande sats är ofta användbar vid beräkningar.

Sats

$$\mathbb{C}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

- ▶ Speciellt gäller följande.

Sats

X och Y är oberoende $\Rightarrow \mathbb{C}(X, Y) = 0$

OBS! Omvändningen är *inte* sann i allmänhet!

- ▶ Exempel?



Korrelation

- ▶ *Korrelationskoefficienten* för X och Y definieras som

$$\rho(X, Y) = \frac{\mathbb{C}(X, Y)}{\mathbb{D}(X)\mathbb{D}(Y)}.$$

- ▶ Notera att korrelationskoefficienten är *dimensionslös*.
- ▶ Man kan visa att $|\mathbb{C}(X, Y)| \leq \mathbb{D}(X)\mathbb{D}(Y)$, vilket medför att $-1 \leq \rho(X, Y) \leq 1$.
- ▶ Man kan även visa att $|\rho(X, Y)| = 1$ om och endast om X och Y är *linjärt beroende*, dvs. $Y = aX + b$ för konstanter a och b ; mer specifikt gäller då att $a > 0 \Rightarrow \rho(X, Y) = 1$ och $a < 0 \Rightarrow \rho(X, Y) = -1$.



Idag

Förra gången

Mer om väntevärden och varianser (Kap. 5.2–5.3)

Beroendemått (Kap. 5.4)

Summor, linjärkombinationer och stora talens lag (Kap. 5.5–5.6)



Varians för summor

- ▶ Vi såg tidigare att

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

för alla (även beroende) s.v. X och Y .

- ▶ I motsvarande uttryck för variansen tillkommer dock en term som motsvaras av kovariansen:

Sats

För alla s.v. X och Y gäller att

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\mathbb{C}(X, Y).$$

- ▶ I specialfallet då X och Y är *oberoende*, vilket implicerar att $\mathbb{C}(X, Y) = 0$, blir således

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y).$$



Varians för linjärkombination

- ▶ Följande sats utvidgar ovan resultat för kovarians för summa av två s.v. till en godtycklig linjärkombination.

Sats

För alla s.v. X_1, \dots, X_n och konstanter a_1, \dots, a_n, b gäller att

$$\mathbb{V} \left(\sum_{i=1}^n a_i X_i + b \right) = \sum_{i=1}^n a_i^2 \mathbb{V}(X_i) + 2 \sum_{1 \leq j < k \leq n} a_j a_k \mathbb{C}(X_j, X_k).$$

- ▶ I specialfallet då alla X_i :na är oberoende blir alla kovarianser noll, och man erhåller

$$\mathbb{V} \left(\sum_{i=1}^n a_i X_i + b \right) = \sum_{i=1}^n a_i^2 \mathbb{V}(X_i).$$

- ▶ OBS! Konstanten b bidrar *ej* till variansen.



Väntevärde och varians för medelvärde

- ▶ Vi låter nu X_1, \dots, X_n vara oberoende s.v. med gemensamt väntevärde μ och standardavvikelse σ . Beteckna *medelvärdet* med

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

- ▶ Med hjälp av ovan satser erhålls

$$\mathbb{E}(\bar{X}) = \frac{1}{n} \mathbb{E} \left(\sum_{i=1}^n X_i \right) = \frac{1}{n} \sum_{i=1}^n \underbrace{\mathbb{E}(X_i)}_{=\mu} = \frac{1}{n} n\mu = \mu$$

och, då X_i :na är oberoende,

$$\mathbb{V}(\bar{X}) = \left(\frac{1}{n} \right)^2 \mathbb{V} \left(\sum_{i=1}^n X_i \right) = \left(\frac{1}{n} \right)^2 \sum_{i=1}^n \underbrace{\mathbb{V}(X_i)}_{=\sigma^2} = \frac{1}{n} \sigma^2.$$



Väntevärde och varians för medelvärde (forts.)

- ▶ Som förväntat gäller att $\mathbb{E}(\bar{X}) = \mu$.
- ▶ Märkvärdigare är att variansen är *omvänt proportionell mot n* enligt

$$\mathbb{V}(\bar{X}) = \frac{1}{n}\sigma^2 \Rightarrow \mathbb{D}(\bar{X}) = \frac{1}{\sqrt{n}}\sigma.$$

- ▶ Det förefaller alltså som om fördelningen för \bar{X} blir mer och mer koncentrerad kring dess väntevärde μ när n ökar.
- ▶ Detta beskrivs närmare av *stora talens lag*.



Stora talens lag

► Sats

Låt X_1, X_2, \dots, X_n vara oberoende och likafördelade s.v. med gemensamt väntevärde μ och låt

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

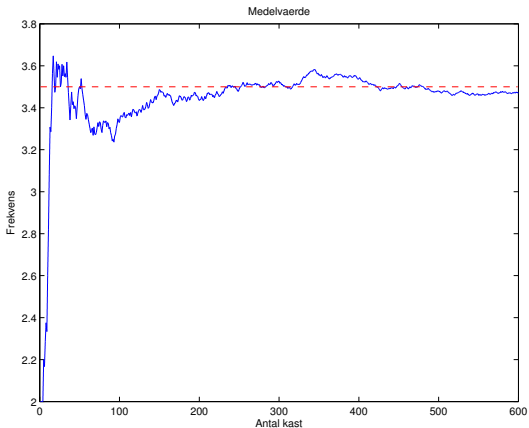
Då gäller, för varje $\varepsilon > 0$,

$$\mathbb{P}(|\bar{X} - \mu| \geq \varepsilon) \rightarrow 0 \quad \text{då} \quad n \rightarrow \infty.$$

► Exempel med tärningskast!



Upprepade tärningskast



Figur: Medelvärde \bar{X} av upprepade tärningskast. Röd streckad linje indikerar väntevärdet $\mu = \mathbb{E}(X) = 3.5$.

Stora talens lag

- ▶ Stora talens lag är ett fundamentalt resultat i matematisk statistik.
- ▶ Tolkning: medelvärdet är en bra uppskattning av väntevärdet när n är stort!
- ▶ Om vi inte känner väntevärdet kan vi alltså gissa, eller "skatta", väntevärdet med hjälp av medelvärdet. Detta kommer att vara av stor vikt när man skall *kalibrera* en slumpmodell med hjälp av observerad data. Stora talens lag är därför en grundsten i empiriska vetenskaper.
- ▶ Detta kommer vi att diskutera vidare i inferensdelen av kursen.



Nästa föreläsning

- ▶ Normalfördelningen,
- ▶ mer om linjärkombinationer av oberoende normalfördelade s.v.,
- ▶ centrala gränsvärdessatsen.

