Föreläsning 7

Bevis av Markovs olikhet

sida 6

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)\,dy = \int_0^a y f_Y(y)\,dy + \int_a^{\infty} y f_Y(y)\,dy \geq$$

$$\geq \int_a^{\infty} y f_Y(y)\,dy \geq a \int_a^{\infty} f_Y(y)\,dy = a\,P(Y \geq a)$$

d.v.s $P(Y \geq a) \leq \frac{E[Y]}{a}$: Markovs olikhet

Bevis av Stora talens lag

$\cancel{P(|\bar{X}_n - \mu|^2 \geq \varepsilon^2)}$

$$P(\mu - \varepsilon < \bar{X}_n < \mu + \varepsilon) =$$

$$= P(|\bar{X}_n - \mu| < \varepsilon)$$

Titta på $P(|\bar{X}_n - \mu|^2 \geq \varepsilon^2) =$

$$[\text{enl Markov}] \leq \frac{E[|\bar{X}_n - \mu|^2]}{\varepsilon^2} = \frac{V(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} =$$

$$= [n \to \infty] \Rightarrow 0$$

$$\Rightarrow P(|\bar{X}_n - \mu| < \varepsilon) \to 1 \quad \text{då } n \to \infty$$

~~Chebyshevs olikhet~~

~~Sannolikheten att ett mätdata hamnar på ett avstånd mer än $h\sigma$ från väntevärdet är~~

$$\cancel{P(|\bar{X} - \mu| > h\sigma) \leq \frac{1}{h^2}}$$

~~bevisas också m.h.a Markov~~

$$\cancel{P\left(\frac{(X-\mu)^2}{\sigma^2} > h^2\right) \leq \left[Y = \frac{X-\mu}{\sigma}\right] \leq \frac{E[Y^2]}{h^2} = \frac{1}{h^2}}$$

Bevis av

<u>Tjebysjeffs olikhet</u>: Sannolikheten att ett mätdata hamnar på ett avstånd $> k\sigma$ från μ är $\leq \frac{1}{k^2}$

$$P\left(|X-\mu| > k\sigma\right) \leq \frac{1}{k^2}$$

$$P\left(|X-\mu| > k\sigma\right) = P\left(\left(\frac{X-\mu}{\sigma}\right)^2 > k^2\right) =$$

$$= \text{Markov} = \leq E\left[\frac{(X-\mu)^2}{\sigma^2}\right] \cdot \frac{1}{k^2} =$$

$$= \frac{V(X)}{V(X)} \cdot \frac{1}{k^2} = \frac{1}{k^2}$$

$$X = \theta + \delta + \varepsilon \leftarrow \text{slumpmässigt fel}$$

X: uppmätt värde

θ: korrekt värde

δ: systematiskt fel

$$E(X) = E[\theta + \delta + \varepsilon] = \theta + \delta + 0 = \theta + \delta = \mu$$

god precision – litet slumpmässigt fel

dålig noggrannhet – stort systematiskt fel

$$\delta = \mu - \varepsilon$$

$$\varepsilon = X - E(X)$$

Sid 1

Normal fördelningen är själv mycket vanlig och mycket kan approximeras till den. Spec C.G.S.

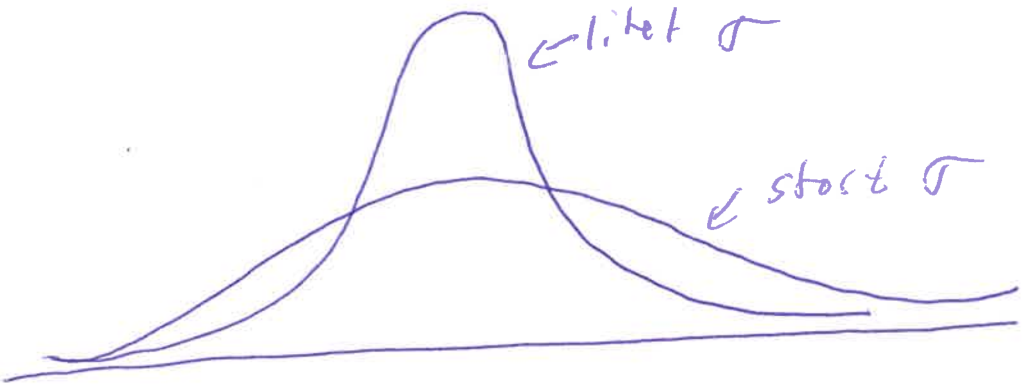
ex Brownsk rörelse där varje tillskott anses vara Normalfördelat

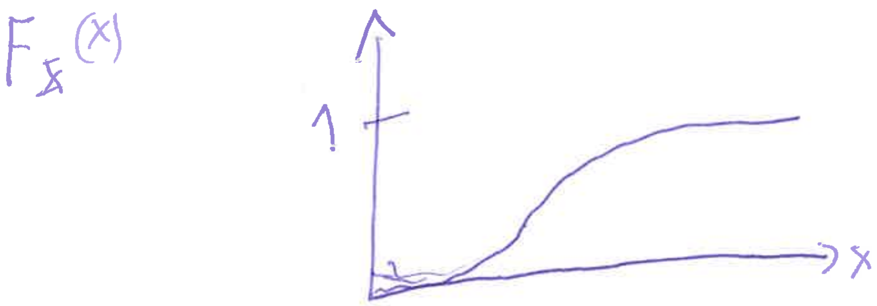
$X \in N(\mu, \sigma)$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

därför använder vi tabell eller räknare





Standardnormalfördelningen har vi när

$X \in N(0,1)$

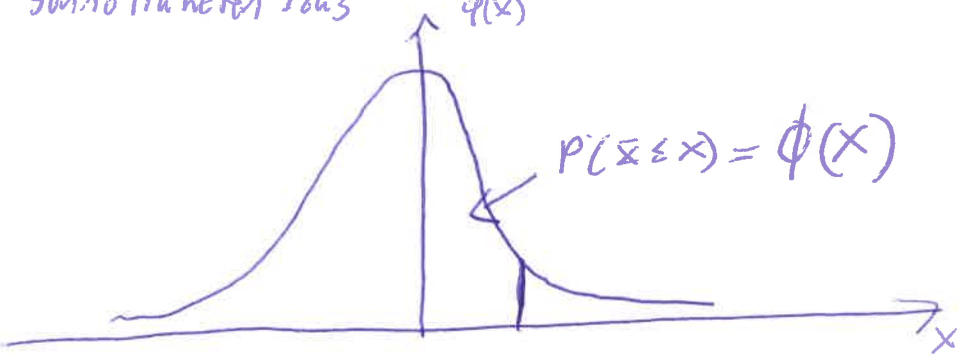
För denna har täthetsfunktionen och fördelningsfunktionen speciella beteckningar nämligen

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

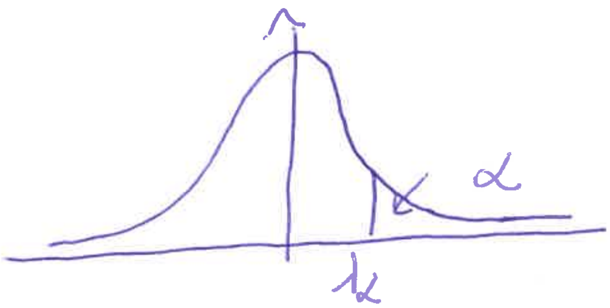
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \, dt$$

Standardnormalfördelningen finns i tab1 och tab2

Tab 1 används när man har en gräns och sannolikheten söks



Tab2 används när man har en sannolikhet och gränsen söks



λ_α kallas α-kvantilen och uppfyller

$P(X > \lambda_\alpha) = \alpha$

Antag att vi har en allmän Normalfördelning

$X \in N(\mu, \sigma)$

Då kan den alltid transformeras till standardnormalfördelningen genom att man bildar $Y = \frac{X-\mu}{\sigma} \Rightarrow Y \in N(0,1)$

Antag $X \in N(\mu, \sigma)$

$$\Rightarrow P(Y \le x) = P\left(\frac{X-\mu}{\sigma} < x\right) =$$

$$= P(X \le \mu + \sigma x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu+\sigma x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

$$= \left[\frac{t-\mu}{\sigma} = u\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du = \Phi(x)$$

$\frac{dt}{\sigma} = du$

$\Rightarrow$ D.v.s $Y \in N(0,1)$

Antag $Y \in N(0,1) \Rightarrow$

$$F_X(x) = P(X \le x) = P\left(\frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right) =$$

$$= P\left(Y < \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$\Rightarrow f_X(x) = \frac{1}{\sigma} \varphi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\frac{x-\mu}{\sigma}\right)^2}{2}}$$

$$= \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma} \varphi\left(\frac{x-\mu}{\sigma}\right)$$

d.v.s $X \in N(\mu, \sigma)$

Räkna ex 6.2 ~~[illegible]~~ sid 152 o 153

$P(\mu-6\sigma < \bar{X} < \mu+6\sigma) = 0.999999998$

$2 \cdot 10^{-9}$

Räkna sedan på sid 149 ~~[illegible]~~

tab 1 $P(\mu-h\sigma < \bar{X} < \mu+h\sigma)$ $h=1,2,3$

$P(\mu-h\sigma < \bar{X} < \mu+h\cdot\sigma) =$ 0.95, 0.99, 0,999 sök h

varje linj-komb av ober N-fördelningar är N-fördelad

Bevisas ej men skrivs opp

även om enskilda mätdata från en försöksserie ej tillhör N-fördelning så gör ofta medelvärdena det om ~~[illegible]~~ antalet mätvärden är stort.

Tag sedan C. G. S. summa och medelvärde

Visa godis

ev ev sats 6.8 sid 158

ex 6.6 sid 160

där a:na och b:na är konstanter och Z_1 och Z_2 är oberoende $N(0,1)$-fördelade s.v.

Man kan slutligen ställa sig frågan varför man inte helt enkelt säger att (X,Y) är tvådimensionellt normalfördelad om X och Y var för sig är normalfördelad. Visserligen vore detta en enkel definition, men klassen av fördelningar skulle bli "för stor" och inte ha önskade egenskaper.

6.7 Centrala gränsvärdessatsen

I detta kapitel har hittills de s.v. antagits vara exakt normalfördelade. Vi skall nu beröra ett av de märkligaste och viktigaste resultaten inom sannolikhetsteorin: Normalfördelningen uppträder i ett generellt sammanhang, som gör den mycket mer tillämpbar än man i förstone kan tro. Man kan nämligen visa att en summa av oberoende likafördelade s.v. med godtycklig fördelning i regel är *ungefär* normalfördelad, bara antalet komponenter i summan är tillräckligt stort.

För att illustrera detta skall vi anknyta till ett tidigare exempel.

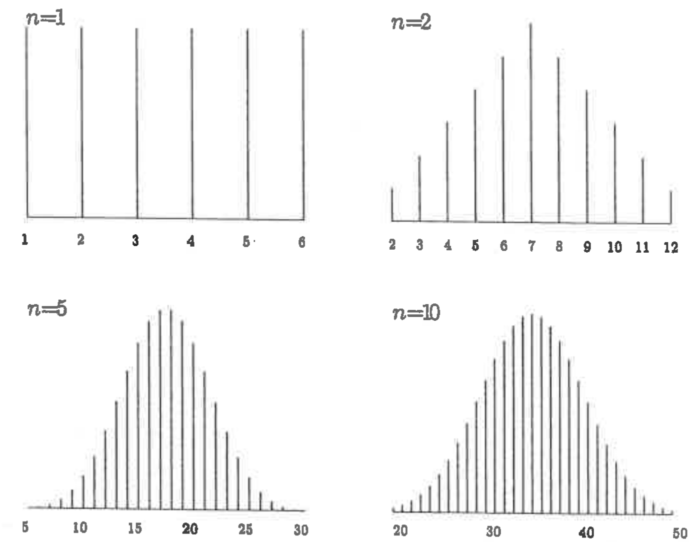
Exempel 6.5 *Poängsumman av tärningskast*
I Exempel 4.9 på sidan 94 betraktades sannolikhetsfunktionen för poängsumman av två oberoende tärningskast. Låt oss nu utsträcka beräkningen till poängsumman av flera tärningskast. Resultatet återges i grafisk form i Figur 6.9, där som jämförelse även ett och två tärningskast medtagits. Man ser att fördelningen snabbt ändras och att den tycks mer och mer få den för normalfördelningen karakteristiska "klockformen". □

Ytterligare en illustration, som visar hur faltning av flera fördelningar ger ungefär normalfördelning, finns i Figur 7.3 på sidan 172. Det sagda är en konsekvens av följande sats.

Sats 6.8 *Centrala gränsvärdessatsen*
Om $X_1, X_2, \ldots$ är en oändlig följd av oberoende likafördelade s.v. med väntevärdet μ och standardavvikelsen $\sigma > 0$, så gäller för $Y_n = X_1 + \cdots + X_n$ att

$$P\Big(a < \frac{Y_n - n\mu}{\sigma\sqrt{n}} \le b\Big) \to \Phi(b) - \Phi(a) \quad \text{då } n \to \infty. \tag{6.18}$$

Lägg märke till att $E(Y_n) = n\mu$, $D(Y_n) = \sigma\sqrt{n}$. För varje givet n är alltså $(Y_n - n\mu)/(\sigma\sqrt{n})$ en standardiserad s.v. (jämför Definition 5.5 på sidan 119). Den har alltså samma väntevärde ($= 0$) och standardavvikelse ($= 1$) som en standardiserad normalfördelad s.v. Satsen



Figur 6.9 *Fördelningen för poängsumman av n tärningskast.*

innehåller ett resultat som är ofantligt mycket skarpare: När n växer mot oändligheten kommer hela fördelningen för den angivna standardiserade variabeln att gå mot en standardiserad normalfördelning.

Beviset av Sats 6.8 är komplicerat och kan inte ges här.

Vi skall införa en definition, som ser komplicerad ut men som visar sig vara praktisk när man arbetar med s.v. som satisfierar centrala gränsvärdessatsen.

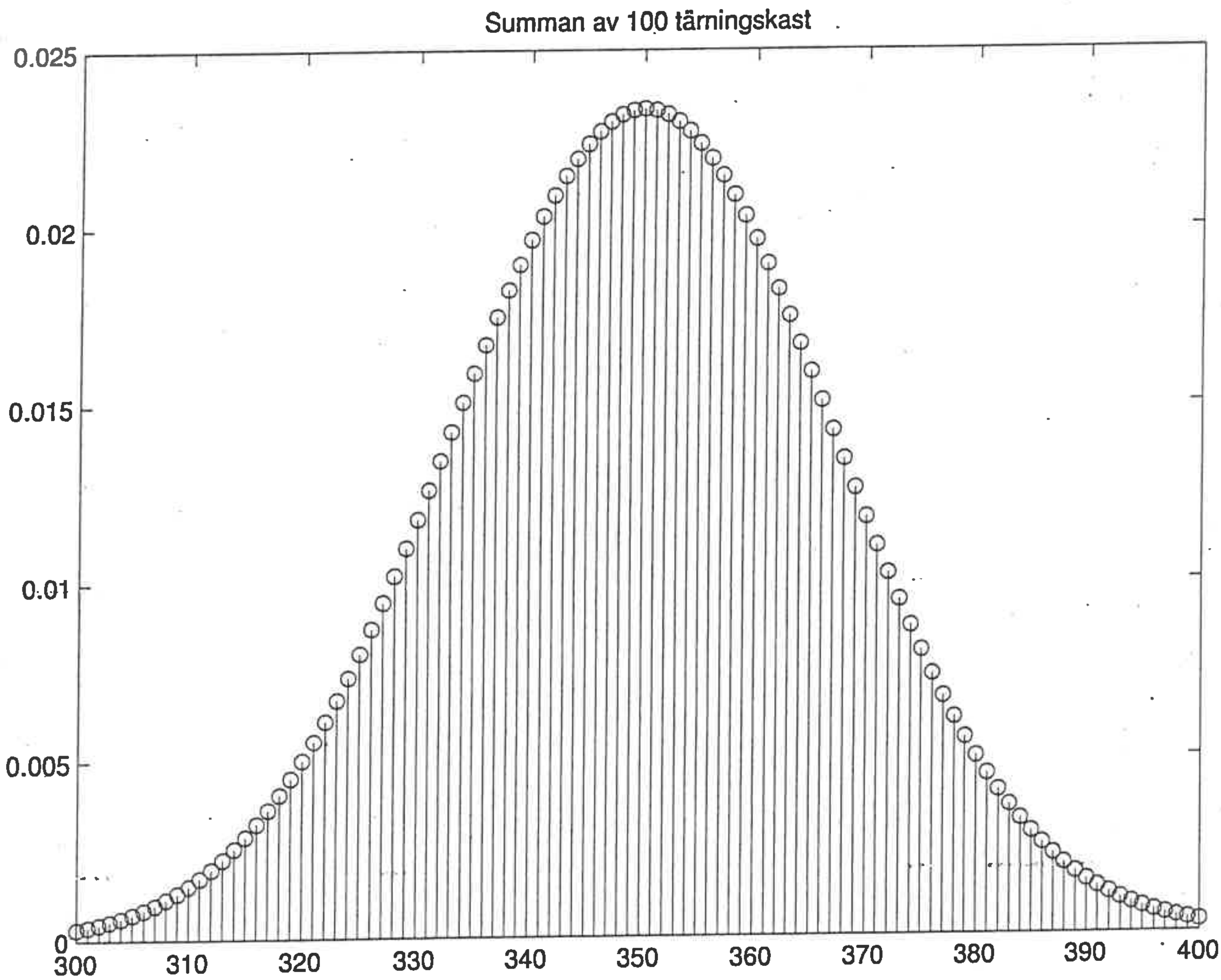
Definition 6.2 Om Z_n, $n = 1, 2, \ldots$, är en oändlig följd av s.v. och man kan finna tal A_n och B_n, $n = 1, 2, \ldots$, sådana att

$$P\Big(a < \frac{Z_n - A_n}{B_n} \le b\Big) \to \Phi(b) - \Phi(a) \quad \text{då } n \to \infty,$$

säges Z_n vara *asymptotiskt normalfördelad* med parametrarna A_n och B_n, eller kortare

$$Z_n \in \mathrm{AsN}(A_n, B_n).$$

Med användning av denna definition utsäger Sats 6.8 att $Y_n \in \mathrm{AsN}(n\mu, \sigma\sqrt{n})$.


Summan av 100 tärningskast
0.025
0.02
0.015
0.01
0.005
0
300
310
320
330
340
350
360
370
380
390
400