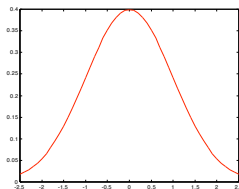


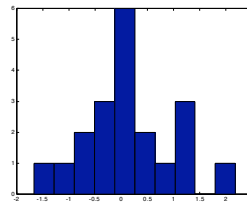
Vi såg i föreläsning ett, att om vi känner den förväntade asymptotiska fördelningen i en given situation så kan vi med utgångspunkt från våra mätdata med hjälp av **minsta kvadrat-metoden** finna vilka parametrar för de asymptotiska fördelningen som bäst motsvarar våra mätdata, gången var:

- 1: Bestäm den förväntade asymptotiska fördelningen. Parametrarna för denna har ett samband med den underliggande verkligheten - det vi vill mäta.
- 2: Gör en mätning av den storhet som representeras av den asymptotiska fördelningen.
- 3: Använd minsta kvadratmetoden för att bestämma vilka parametrar för den asymptotiska fördelningen som ger bäst överensstämmelse mellan teori och experiment.

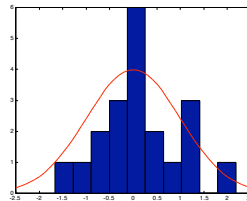
Teorin ger oss en asymptotisk fördelning



Mätningar ger oss en verklig fördelning



Minsta kvadratmetoden låter oss bestämma vilka värden på de teoretiska parametrarna som ger bästa överensstämmelsen



Nästa fråga blir då:

- hur skulle vi kunna definiera en asymptotisk fördelning för till exempel mätningen av längden av en parkeringsficka?

Svaret ges av en vid första påseende märklig sats:

Centrala Gränsvärdessatsen:

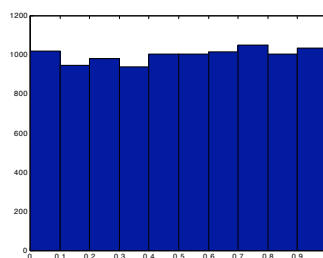
Om vi summerar ett stort antal slumpmässigt fördelade tal, så kommer den asymptotiska fördelningen för summan att gå mot en normalfördelning

Detta gäller oberoende av hur fördelningen ser ut för de termer som ingår i summan!!

Exempel:

Summera tal slumpvis fördelade mellan 0 och 1:

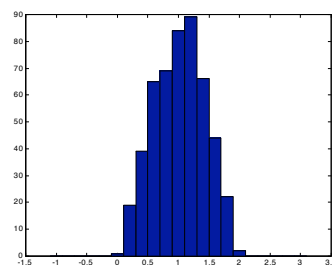
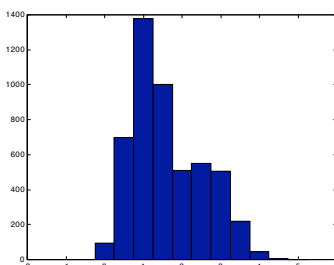
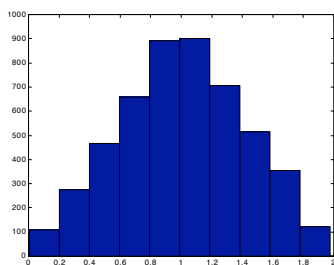
Den ursprungliga fördelningen:



Summan av 2 tal:

5 termer

20 termer



För att förstå hur denna sats kan hjälpa oss så kan vi betrakta resultatet av en mätning som beroende av det sanna värdet till vilket har adderats slumpvisa bidrag från ett stort antal okända (och en del kända) felkällor. Om vi utgår från att det inte finns några systematiska effekter kommer positiva fel vara lika vanliga som negativa. Resultatet av ett stort antal mätningar kommer då att spridas runt det sanna värdet, och fördelningen av mätningarna runt detta kommer att ha en form som ges av normalfördelningen.

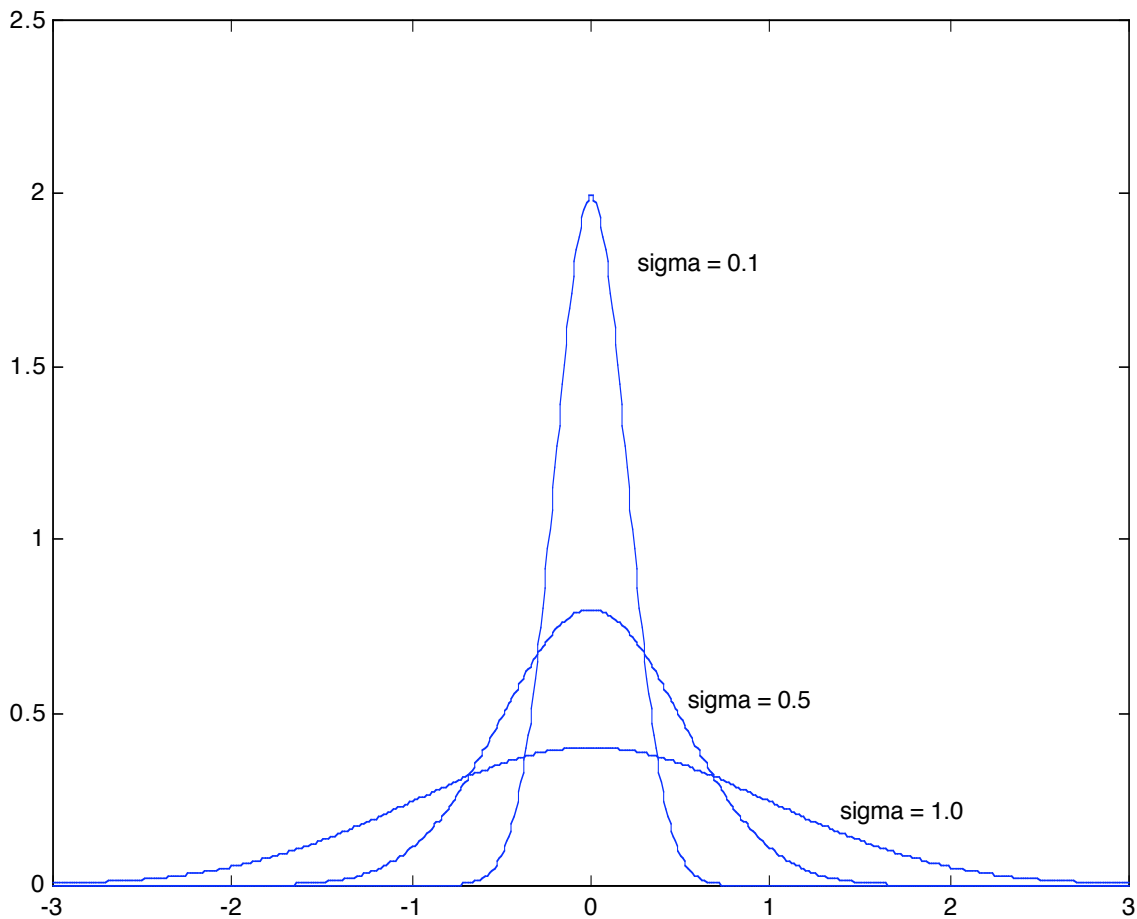
Normalfördelningen:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\}$$

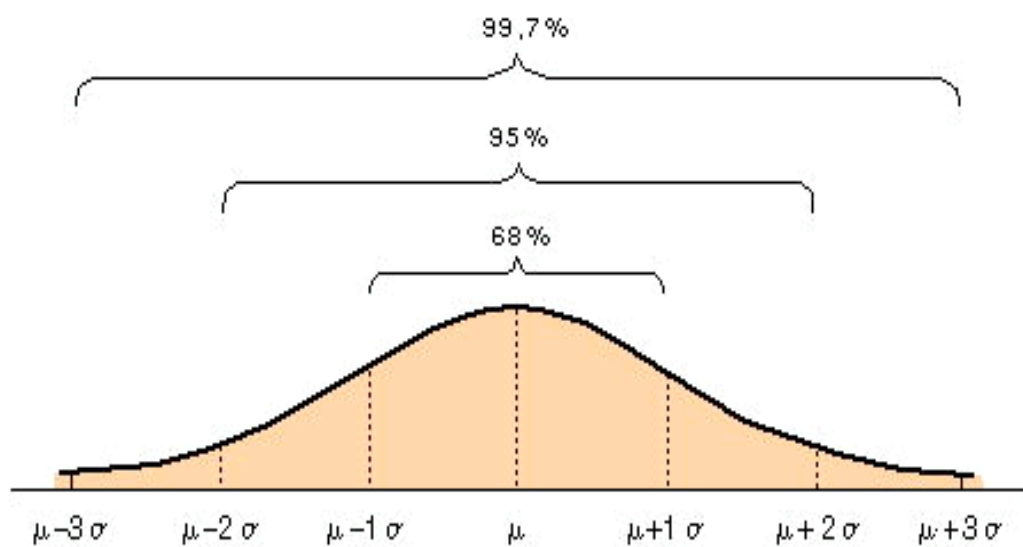
Observera att denna fördelning är normaliserad till ett.

Egenskaper

- Maximum vid $x = \mu$
- Symmetrisk runt $x = \mu$
- När σ är litet så blir exponenten stor $\rightarrow$ lutningen blir större
- När σ är litet så blir normaliseringskonstanten större $\rightarrow$ höjden vid toppen blir relativt sett högre.



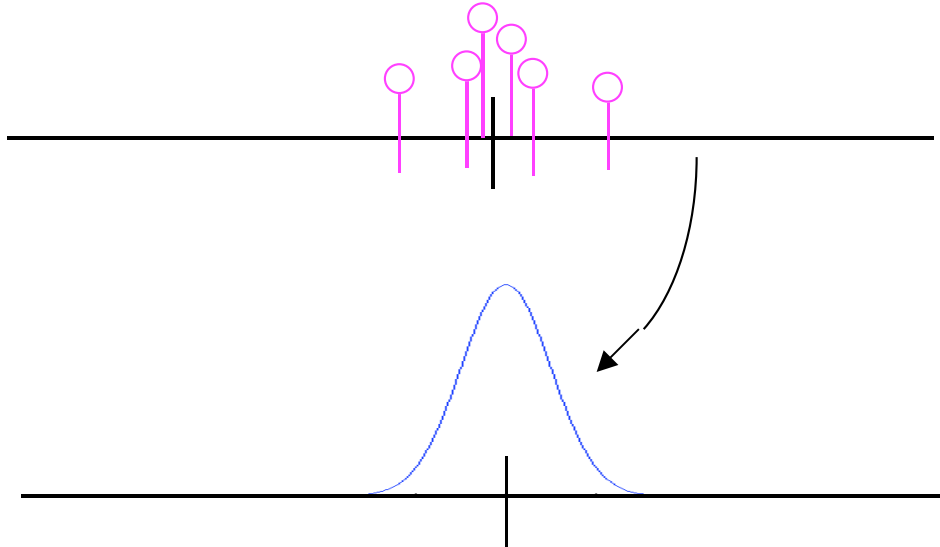
Normalfördelningen har normaliseringen 1. Uttrycker man integrationsgränser i parametern σ så har alla normalfördelningar samma area inom dessa gränser - oberoende av vilka exakta värden μ och σ antar. Integrerar man en normalfördelning mellan till exempel $-\sigma$ och $+\sigma$ så är arean 68% av hela arean. Detta har betydelse när vi tolkar $f(x)$ som en sannolikhetsfördelning - sannolikheten att hamna i intervallet $[\mu - 2\sigma, \mu + 2\sigma]$ är 95% och så vidare.



Tillbaks till vår mätsituation:

Om det inte finns stora systematiska effekter så kan vi alltså förvänta oss att våra mätresultat - efter ett stort antal mätningar och under förutsättning att det inte finns systematiska effekter - beskrivs av en normalfördelning.

Vi förväntar oss att $\bar{x}$ svarar mot det sanna värdet för den parameter vi vill mäta, vi kommer snart att se att s säger oss någonting om mätmetodens precision.



Givet mätdata, hur uppskattar vi parametrarna μ och σ ?

Lösningen ligger i att vi betraktar normalfördelningen som en sannolikhetsdistribution för att göra observationen x , givet μ och σ . Om den underliggande distributionen är en normalfördelning med parametrarna μ och σ , så är sannolikheten för att göra observationen x proportionell mot:

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

Oberoende sannolikheter:

för oberoende händelser (sannolikheten av en händelse påverkas inte av utfallet i de andra händelserna) är sannolikheten för serien lika med produkten av sannolikheterna för de enskilda händelserna.



Om våra mätningar av x är oberoende, vilket vi kommer att anta, så ges tydligen sannolikheten för att observera just vår mätserie, $x_1, x_2, x_3, \dots, x_N$ av produkten av sannolikheterna för de enskilda mätningarna:

$$P(x_1, x_2, x_3, \dots, x_N) = P(x_1) \cdot P(x_2) \cdot P(x_3) \cdot \dots \cdot P(x_N)$$

$$P(\text{serien}) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{(x_1 - \mu)^2}{2\sigma^2}\right] \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{(x_2 - \mu)^2}{2\sigma^2}\right] \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{(x_3 - \mu)^2}{2\sigma^2}\right] \cdots$$

$$\cdots \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{(x_N - \mu)^2}{2\sigma^2}\right]$$

Uttrycket ger oss sannolikheten att observera vår mätserie $x_1, x_2, x_3, \dots, x_N$ givet parametrarna μ och σ för den underliggande fördelningen. Vi bestämmer nu vilka värden på μ och σ som ger oss den största sannolikheten att observera just vår mätserie genom normal maximering:

$$\frac{\partial P}{\partial \mu} = \sum_{i=1}^N \frac{\partial}{\partial \mu} \left(\frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right] \right) = 0$$

Villkoret att denna derivata är noll ger oss då ett estimat för μ :

$$\frac{\partial P}{\partial \mu} = 0 \quad \Rightarrow \quad \sum_{i=1}^N (x_i - \mu) = 0 \quad \Rightarrow \quad \hat{\mu} = \frac{1}{N} \sum_{i=1}^N x_i$$

Den bästa uppskattningen av parametern σ ges alltså av medelvärdet för alla ingående storheter.

Vi gör nu samma optimering för parametern σ :

$$\frac{\partial P}{\partial \sigma} = \frac{\partial}{\partial \sigma} \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) \cdot \sum_{i=1}^N \exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right] + \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \sum_{i=1}^N \frac{\partial}{\partial \sigma} \left(\exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right] \right)$$

$$= \frac{\partial}{\partial \sigma} \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) \cdot \sum_{i=1}^N \exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right] + \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \sum_{i=1}^N \left(-\frac{(x_i - \mu)^2}{\sigma^3} \right) \cdot \exp\left[-\frac{(x_i - \mu)^2}{2\sigma^2}\right]$$

$$\frac{\partial P}{\partial \sigma} = 0 \quad \Rightarrow \quad \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

I detta uttryck betecknar μ det sanna värdet på denna parameter. Eftersom detta värde oftast är okänt så får vi approximera det med vårt bästa estimat - medelvärdet. Man kan visa (det görs inte i kursen) att detta medför att vi får ett något modifierat uttryck för estimatet:

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 = s^2$$

Även om vi inte strikt visar att uttrycket modifieras på det här sättet så kan vi intuitivt förstå att det är riktigt:

- För en enda mätt punkt kan vi inte ha någon uppfattning om spridningen (om vi inte vet runt vilket värde denna enda punkt är spridd). Rent matematiskt motsvarar uttrycket denna förväntning eftersom s^2 i detta fall blir av typen $0/0$, alltså ej definierat.

När vi har mätt N punkter finns det N "frihetsgrader" (vi återkommer senare till detta begrepp), det vill säga det finns N oberoende "fakta" att tillgå. Men när vi ersätter $\bar{x}$ med medelvärdet så använder vi data till att beräkna medelvärdet, vilket gör att vi bara har $N-1$ frihetsgrader kvar. I princip kan vi varje gång x_1 dyker upp ersätta det med

$$x_1 = N \cdot \bar{x} - \sum_{i=2}^N x_i$$

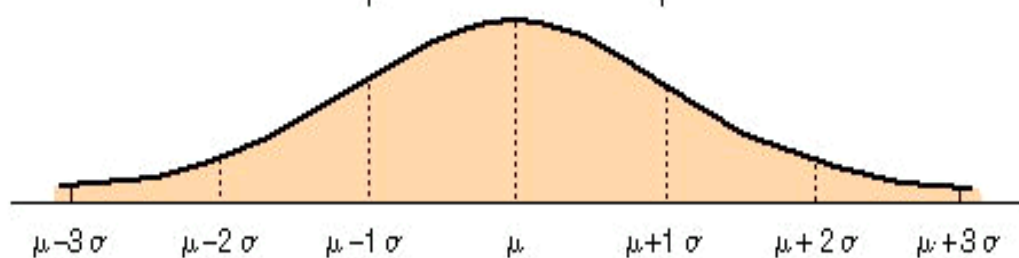
Vilket demonstrerar att det bara finns $N-1$ oberoende variabler kvar.

Så vad betyder σ ?

Om vi tolkar normalfördelningen som den asymptotiska fördelningen för en mätning så betyder det att sannolikheten för att vår mätning skall ge ett resultat i intervallet $[x, x+dx]$ lika med arean under kurvan mellan dessa värden eller:

$$P(x \in [x, x+dx]) = \int_x^{x+dx} f(y) dy$$

Där $f(y)$ betecknar normalfördelningen centrerad kring μ



Gör vi en mätning är sannolikheten att den hamnar inom ± 1 sigma från det sanna värdet 68%. Så σ , eller estimatet för σ , säger hur stor spridningen kring det sanna värdet, det är en **egenskap hos mätmetoden**.

Några ord på vägen nu när ni skall börja mäta:

Tag det lugnt!

Fem minuter tjänade på övningslab kostar ofta en halvtimme i datasalen när man försöker reda ut vad man egentligen gjorde däruppe.

Förbered er

Väl uppe på övningslabb, med labbkompisen flåsande i nacken är det för sent att läsa in vad man skall göra.

Släng ingenting! Inga lösa lappar!

Det finns hur många exempel som helst på att det i efterhand dyker upp frågor som man inte tänkt på. Inte ens när laborationen är redovisad och godkänd kan man vara säker på att man inte i något annat sammanhang vill gå tillbaks till det man gjort för att förstå någon aspekt av det.

Av samma skäl skall allt föras in i logboken, inte skrivas upp på lösa lappar som sedan ofrånkomligt försvinner.

Dokumentera upptställningen och yttre variabler!

Era anteckningar skall vara så fullständiga att ni kan komma tillbaks och göra om identiskt samma försök - samma mätinstrument, samma fjädrar, resistorer mm mm.

Detta kan vara enda sättet att undersöka resultat som i efterhand ser konstiga ut, om någon ifrågasätter era resultat måste ni kunna säga antingen att ni kan eller att ni inte kan reproducera dem. Ni får inte hamna i en situation där ni inte vet om ni tror på ert resultat eller inte.

Dokumentera under arbetets gång!

På samma sätt som det är viktigt att dokumentera de yttre förutsättningarna så är det viktigt att dokumentera arbetets gång. I vilken ordning gjordes mätningarna. Ofta kan man t ex misstänka att mätinstrument "drivit" under arbetets gång. Har man koll på i vilken ordning man gjort vad är det enkelt att kontrollera.

Beräkna delresultat!

Listan på laborationsrapporter som mynnar ut i "de här resultaten verkar helt orimliga men nu i efterhand kan jag inte förstå vad som gick fel..." är lång. Se till att din rapport inte hamnar där! Grova överslagsberäkningar och enkla diagram gjorda under laborationens gång kan ofta avslöja om det smugit sig in någon katt bland hermelinerna.

Vänd om!

Vi kommer att bereda er möjligheten att titta upp på labbet "dagen efter" om ni behöver kolla något. Under "självständigt arbete" på schemat är labb- och datorsalar bokade för kursen och assistenterna finns i huset.

Misstänker man att något gått galet bör man gå tillbaks och göra en kontrollmätning