Matematisk statistik för B, K, N, BME och Kemister

Johan Lindström

Repetition

Grundläggande begrepp (Kap. 3.1)

- **Utfall** – resultatet av ett slumpmässigt försök.
  Bet. $\omega_1, \omega_2, \ldots$
- **Händelse** – en samling av ett eller flera utfall.
  Bet. $A, B, \ldots$
- **Utfallsrum** – mängden av möjliga utfall.
  Bet Ω

Oberoende händelser (Kap. 3.2.4)

Händelserna A och B är **oberoende** av varandra

$$\iff$$

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

Stokastisk variabel (Kap. 3.3)

En **stokastisk variabel** eller **slumpvariabel** är ett **tal** vars värde styrs av slumpen.
Bet $X, Y, \ldots$.

En stokastisk variabel beskrivs av:

Sannolikhetsfunktion För en diskret s.v. X

$$p_X(k) = \mathrm{P}(X = k)$$

Täthetsfunktion För en kontinuerlig s.v X har vi $f_X(x)$.

$$\mathrm{P}(a \leq X \leq b) = \int_a^b f_X(x)\, dx$$

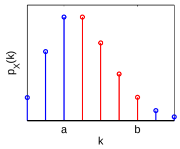
Fördelningsfunktion Summa av $p_X(k)$ eller integral av $f_X(x)$.

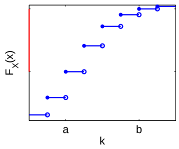
$$F_X(x) = \mathrm{P}(X \leq x)$$

Fördelningsfunktioner (Kap. 3.3)

Diskret

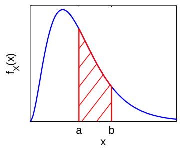
$$P(a < X \le b) = \sum_{k=a+1}^{b} p_X(k) \qquad P(a < X \le b) = F_X(b) - F_X(a)$$

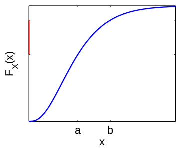




Kontinuerligt

$$P(a < X \le b) = \int_a^b f_X(x)\,dx \qquad P(a < X \le b) = F_X(b) - F_X(a)$$





Standardfördelningar (Kap. 3.6 & 6)

Diskret fördelning:

Binomialfördelning Ett slumpmässigt försök som lyckas med slh. p upprepas n oberoende ggr, $X =$ Antal ggr försöket lyckas.

Poissonfördelning Räknar antal händelser.

Kontinuerlig fördelning:

Rektangel- eller likformig fördelning Lika fördelade händelser i intervall.

Exponentialfördelning Ofta överlevnadstid, eller tid till/mellan händelser.

Normalfördelning Summor av **många** oberoende, vanligt antagande om för mätfel.

Väntevärde, $E(X), \mu, \mu_X, m, \ldots$ (Kap. 3.5)

Väntevärdet anger **tyngdpunkten** för fördelningen och kan tolkas som det värde man får i "**medeltal** i långa loppet".

$$E(X) = \begin{cases} \int_{-\infty}^{\infty} x f_X(x)\,dx & \text{Kont.} \\ \sum_k k p_X(k) & \text{Diskr.} \end{cases}$$

Varians, $V(X), \sigma^2, \sigma_X^2$ (Kap. 3.5)

Variansen anger hur utspridd X är kring sitt väntevärde.

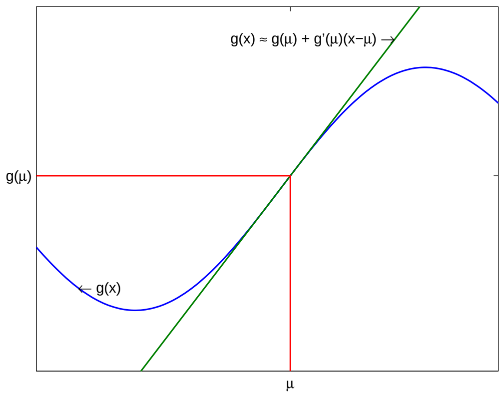
$$V(X) = E\left\{\left[X - E(X)\right]^2\right\} = E(X^2) - E(X)^2$$

Standardavvikelse:, $D(X), \sigma, \sigma_X$ $\qquad D(X) = \sqrt{V(X)}$

Räkneregler för Väntevärde och Varians (Kap. 3.5.4 & 4.4)

- $\mathrm{E}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i \mathrm{E}(X_i)$
- $\mathrm{V}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \mathrm{V}(X_i)$ om oberoende

Linjärisering av $g(x)$ kring punkten $\mu = E(X)$

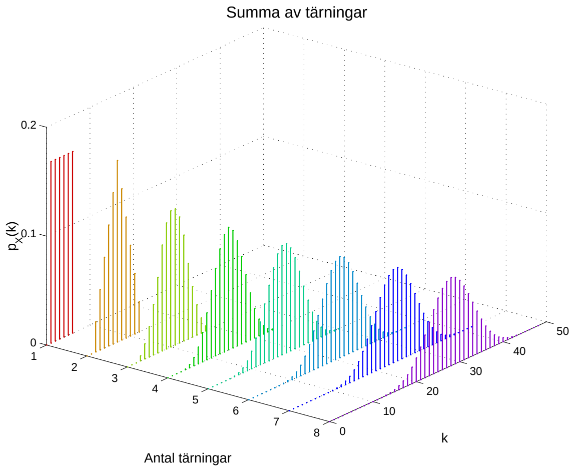


Gauss approximationsformler i en variabel (Kap. 5.2)

$Y = g(X)$. Taylorutveckla funktionen g kring $\mu = E(X)$

$$g(X) \approx g(\mu) + (X - \mu)g'(\mu) \quad \Longrightarrow$$

- $E(Y) \approx g(E(X))$
- $V(Y) \approx g'[E(X)]^2 V(X)$



Centrala gränsvärdessatsen CGS (Kap. 4.5)

Om $X_1, X_2, \ldots, X_n$ är oberoende likafördelade stokastiska variabler med $E(X_i) = \mu$, $V(X_i) = \sigma^2$ så är

$$\sum_{i=1}^{n} X_i \underset{\sim}{\in} \mathsf{N}\left(n\mu, n\sigma^2\right) \qquad \text{då } n \text{ stort} \quad (n \to \infty)$$

1. Om $Y = \sum_{i=1}^{n} X_i$ gäller $Y \underset{\sim}{\in} \mathsf{N}\left(n\mu, n\sigma^2\right)$
2. Om $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ gäller $\overline{X}_n \underset{\sim}{\in} \mathsf{N}\left(\mu, \frac{\sigma^2}{n}\right)$

Statistikteori – översikt

Punktskattning
Hur gör man en bra gissning av en okänd storhet? Hur vet man att den är bra?

Intervallskattning
Hitta istället ett intervall som täcker den okända storheten med en given (stor) sannolikhet.

Hypotestest
Om gissningen blev 0.013, kan rätt värde på den okända storheten ändå vara 0.01?

Regression
Hur vet vi om två variabler påverkar varandra?

Försöksplanering & Faktorförsök
Hur konstruerar man studier som på bäst sätt (minst antal mätningar) undersöker effekten av olika faktorer (behandlingar)?

Statistikteori, grundläggande begrepp (Kap. 7.1)

Stickprov

Ett **stickprov**, $x_1, x_2, \ldots, x_n$, är **observationer** av s.v. $X_1, \ldots, X_n$ från någon fördelning $X_i \in F(\vartheta)$ där ϑ är en okänd **parameter**.

Skattning

En **skattning** av ϑ, $\vartheta^*(x_1, \ldots, x_n)$ är en observation av den s.v. $\vartheta^*(X_1, \ldots, X_n)$. Båda betecknas oftast bara med ϑ^*.

Bra egenskaper för en skattning är

Väntevärdesriktig: $E(\vartheta^*) = \vartheta$, inget systematiskt fel.

Effektiv: liten varians (osäkerhet) $V(\vartheta^*)$.

Konfidensintervall (Kap. 7.3 & 9)

Ett **konfidensintervall** för en parameter ϑ täcker rätt värde på ϑ med sannolikheten $1 - \alpha$.

$1 - \alpha$ kallas **konfidensgrad**. Vanliga värden är 0.95, 0.99 och 0.999.

Normalfördelad skattning, $\vartheta^* \in \mathsf{N}\left(\vartheta, \mathsf{V}(\vartheta^*)\right)$

$\mathsf{D}(\vartheta^*)$ **känd:** $I_\vartheta = \vartheta^* \pm \lambda_{\alpha/2}\mathsf{D}(\vartheta^*)$

$\mathsf{D}(\vartheta^*)$ **okänd:** $I_\vartheta = \vartheta^* \pm t_{\alpha/2}(f)d(\vartheta^*)$

Normalapproximation, $\vartheta^* \underset{\sim}{\in} \mathsf{N}\left(\vartheta, \mathsf{V}(\vartheta^*)\right)$ (Ex: CGS)

$\mathsf{D}(\vartheta^*)$ **känd:** $I_\vartheta = \vartheta^* \pm \lambda_{\alpha/2}\mathsf{D}(\vartheta^*)$

$\mathsf{D}(\vartheta^*)$ **okänd:** $I_\vartheta = \vartheta^* \pm \lambda_{\alpha/2}d(\vartheta^*)$ (alltid λ-kvantil)

Konfidensintervall för σ^2 i $\mathsf{N}\left(\mu, \sigma^2\right)$ (Kap. 8.1)

$x_1, \ldots, x_n$ observationer av $X_i \in \mathsf{N}\left(\mu, \sigma^2\right)$

Ett $1 - \alpha$ konfidensintervall för σ^2 ges av

$$I_{\sigma^2} = \left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)} \right)$$

I allmänhet har vi

$$I_{\sigma^2} = \left(\frac{f \cdot s^2}{\chi^2_{\alpha/2}(f)}, \frac{f \cdot s^2}{\chi^2_{1-\alpha/2}(f)} \right)$$

där f är antalet frihetsgrader

Sammanvägd variansskattning (Kap. 7.4 & 7.7)

Om vi har

$$x_1, \ldots, x_{n_x} \qquad \text{obs. av } X_i \in \mathrm{N}\left(\mu_x, \sigma^2\right)$$
$$y_1, \ldots, y_{n_y} \qquad \text{obs. av } Y_i \in \mathrm{N}\left(\mu_y, \sigma^2\right)$$

kan den gemensamma variansen σ^2 skattas med

$$s_p^2 = \frac{(n_x-1)s_x^2 + (n_y-1)s_y^2}{n_x - 1 + n_y - 1} = \frac{Q}{f}, \qquad \left(\frac{Q}{\sigma^2} \in \chi^2(f)\right)$$

Ett konfidensintervall för $\mu_x - \mu_y$ blir t.ex.

$$I_{\mu_x - \mu_y} = \bar{x} - \bar{y} \pm t_{\alpha/2}(f)\, s_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$$

eftersom $\mu_x^* - \mu_y^* = \bar{X} - \bar{Y} \in \mathrm{N}\left(\mu_x - \mu_y, \sigma^2\left(\frac{1}{n_x} + \frac{1}{n_y}\right)\right)$.

Stickprov i par (Kap. 7.8)

Vid många mätsituationer är det vanligt att man mäter före och efter en "behandling" på n inbördes olika föremål.

Modell:
Före: $X_i \in \mathrm{N}\left(\mu_i, \sigma_1^2\right)$ Efter: $Y_i \in \mathrm{N}\left(\mu_i + \Delta, \sigma_2^2\right)$

Bilda $Z_i = Y_i - X_i \in \mathrm{N}\left(\Delta, \sigma^2\right)$ och skatta Δ med $\bar{z}$.
Gör konfidensintervall som vanligt för ett stickprov, dvs

$$I_\Delta = \bar{z} \pm t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}},$$
$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z})^2.$$

Hypotesprövning (Kap. 7.5 & 9)

H_0 **förkastas** om observationerna, ϑ^*, avviker för mycket från **nollhypotesen** ϑ_0.

Testa **nollhypotesen**	$H_0:\ \vartheta = \vartheta_0$
mot **mothypotesen** (tex)	$H_1:\ \vartheta \neq \vartheta_0$

på nivån α; **felrisken** α ges av

$$\alpha = \mathrm{P}(H_0 \text{ förkastas trots att den är sann})$$

De vanligaste mothypoteserna är

- $H_1: \vartheta \neq \vartheta_0$ H_0 förkastas om ϑ^* avviker för långt från ϑ_0 både uppåt och nedåt.
- $H_1: \vartheta < \vartheta_0$ H_0 förkastas om ϑ^* är tillräckligt mycket $< \vartheta_0$.
- $H_1: \vartheta > \vartheta_0$ H_0 förkastas om ϑ^* är tillräckligt mycket $> \vartheta_0$.

Olika metoder för att utföra hypotestest

1. **Direktmetoden eller P-värde** **(Def. 7.35)**
   - Antag att H_0 är sann
   - Räkna ut **P-värdet** $p = \mathrm{P}(\text{Få det vi fått eller värre})$
   - Om $p < \alpha$ förkastas H_0
2. **Konfidensmetoden** **(Kap. 7.5.2)**
   Gör ett konfidensintervall med konfidensgraden $1-\alpha$ och förkasta H_0 på nivån α om intervallet ej täcker ϑ_0.
   Intervallen skall, beroende på H_1, vara

Test	$H_1: \vartheta < \vartheta_0$	$H_1: \vartheta \neq \vartheta_0$	$H_1: \vartheta > \vartheta_0$
Intervall:	uppåt begr	tvåsidigt	nedåt begr

3. **Testkvantitet $T(X)$ och kritiskt område C** **(Kap. 7.5.3)**
   Förkasta H_0 om testskvantiteten hamnar i det kritiska området.
   C och T skall väljas så att
   $$\alpha = P(T(X) \in C) = P(\text{"Förkasta } H_0 \text{ om } H_0 \text{ är sann"})$$

Hypotestest – Vilken metod?

- Normalfördelad skattning.
  - σ känd: Vilken som helst.
  - σ okänd: Direktmetoden kräver t-fördelningens fördelningsfunktion.
- Fördelning där $\mu^* = \bar{X} \underset{\sim}{\in} \mathrm{N}(\mu, \mathrm{V}(\mu^*))$ ... enl. CGS.
  - Vilken som helst
- Bin, Po, ... där $\mathrm{D}(\vartheta^*)$ innehåller ϑ.
  - Direktmetoden: **Går alltid** att använda, ibland med normalapproximation.
  - Testkvantitet: Kräver normalt **normalapproximation**.

Vid **styrkefunktion** är det naturligt att **utgå från testkvantitet**.

Testkvantiter

Antag att vi vill testa $H_0: \ \vartheta = \vartheta_0$.

Model		Skattning	$T(X)$	$D(\vartheta^*)/d(\vartheta^*)$	kvantil
$X_i \in \mathrm{N}(\mu, \sigma^2)$	σ känd	$\mu^* = \bar{X}$	$\frac{\mu^*-\mu_0}{\mathrm{D}(\mu^*)}$	$\frac{\sigma}{\sqrt{n}}$	λ
	σ okänd		$\frac{\mu^*-\mu_0}{d(\mu^*)}$	$\frac{s}{\sqrt{n}}$	$t(f)$
$X \in \mathrm{Bin}(n,p)$		$p^* = \frac{x}{n}$	$\frac{p^*-p_0}{\mathrm{D}_0(p^*)}$	$\sqrt{\frac{p_0(1-p_0)}{n}}$	λ
$X_i \in \mathrm{Po}(\mu)$		$\mu^* = \bar{X}$	$\frac{\mu^*-\mu_0}{\mathrm{D}_0(\mu^*)}$	$\sqrt{\frac{\mu_0}{n}}$	λ

Notera:

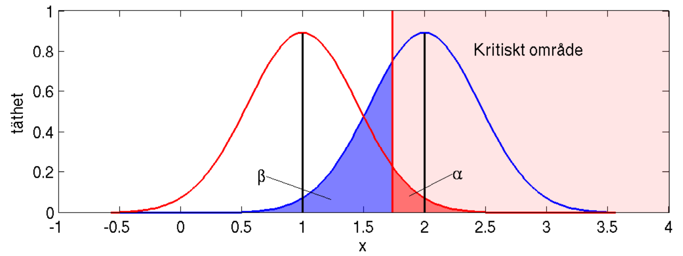
1. Standardavvikelse/medelfel **räknas under H_0**.
2. Bin och Po fallet kräver **normalapproximation**.
3. α-kvantil om **ensidigt**, $\alpha/2$-kvantil om **tvåsidigt**.

Styrkefunktion (Kap. 7.6)

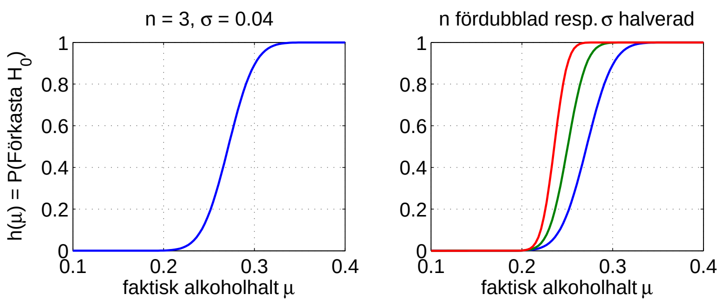
Användas för att avgöra hur bra testet skiljer H_0 från H_1.

$$h(\vartheta) = \text{P}(\text{"Förkasta } H_0 \text{ om } \vartheta \text{ är rätt värde"})$$

Typ 1 fel: $\alpha = P(H_0 \text{ förkastas om } H_0 \text{ sann})$
Typ 2 fel: $\beta = P(H_0 \text{ förkastas } \textbf{ej} \text{ om } H_0 \textbf{ ej} \text{ sann})$



Styrkefunktion för testet av promillehalt ($H_0 : \mu = 0.2$)



Den okända sanningen	Nykter	Olovligt påverkad	
Mätresultat $\mu^* = \bar{x}$		Säkerhetsmarginal	Kritiskt område
Slutsats från test	Frikänns		Döms
	0	0.2	0.27

μ

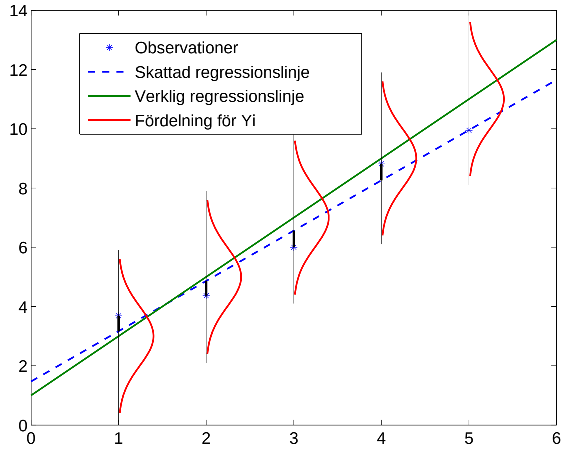
Linjär regression

Modell (Kap. 10.2)

Vi har n st par av mätvärden (x_i, y_i), $i = 1, \ldots, n$ där y_i är observationer av

$$Y_i = \alpha + \beta x_i + \varepsilon_i$$

där ε_i är oberoende av varandra, och $\varepsilon_i \in \text{N}\left(0, \sigma^2\right)$.



Parameterskattningarna (Kap. 10.4–10.5)

Skattningarna av α^*, β^*

$$\beta^* = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}} \in \mathrm{N}\left(\beta, \frac{\sigma^2}{S_{xx}}\right)$$

$$\alpha^* = \bar{y} - \beta^* \bar{x} \in \mathrm{N}\left(\alpha, \sigma^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)\right)$$

och $s^2 = (\sigma^2)^*$ är

$$s^2 = \frac{Q_0}{n-2} \text{ där } Q_0 = \sum_{i=1}^n (y_i - \alpha^* - \beta^* x_i)^2 = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$$

$$\frac{Q_0}{\sigma^2} \in \chi^2(n-2)$$

Skattningarna α^* och β^* är dock **inte oberoende** av varandra.

Konfidens-, prediktions- och kalibreringsintervall:

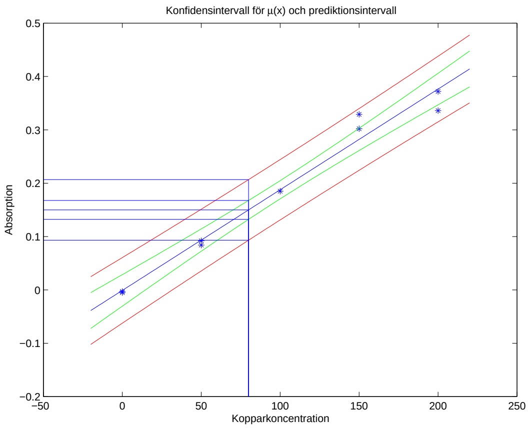
$$I_\beta = \beta^* \pm t_{a/2}(n-2)\frac{s}{\sqrt{S_{xx}}} \qquad I_\alpha = \alpha^* \pm t_{a/2}(n-2)s\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}}$$

$$I_{\mu_0} = \alpha^* + \beta^* x_0 \pm t_{a/2}(n-2)s\sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}.$$

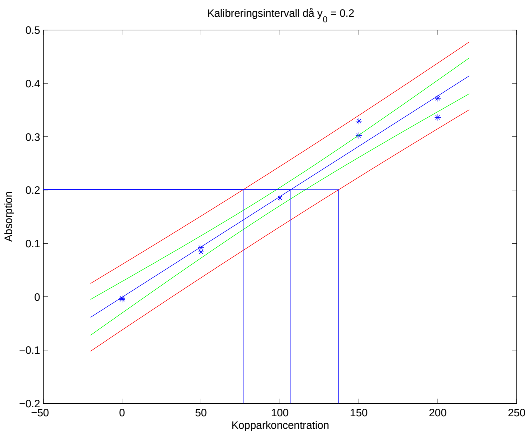
$$I_{Y(x_0)} = \alpha^* + \beta^* x_0 \pm t_{a/2}(n-2)s\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

$$I_{x_0} = x_0^* \pm t_{a/2}(n-2) \cdot \frac{s}{|\beta^*|} \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_0^* - \bar{x})^2}{S_{xx}}}$$

Konfidens- och prediktionsintervall



Kalibreringsintervall



Residualanalys/Modellvalidering (Kap. 10.10)

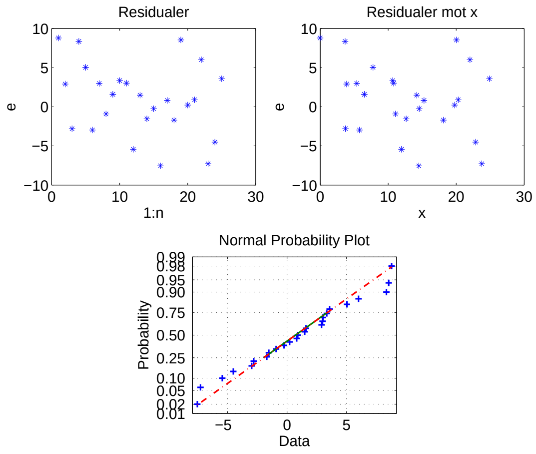
För att undersöka hur bra modellen stämmer kan vi kan studera **residualerna**, dvs avvikelserna mellan observerade y-värden och den skattade linjen.

$$e_i = y_i - \alpha^* - \beta^* x_i, \quad i = 1, \dots, n$$

Dessa är observationer av ε_i, och residualerna bör alltså:

- se ut att komma från en och samma **normalfördelning**
- vara **oberoende** av varandra
- vara **oberoende** av alla x_i.

Bra residualplottar



Multipel linjär regression (Kap. 11.2)

Modellen

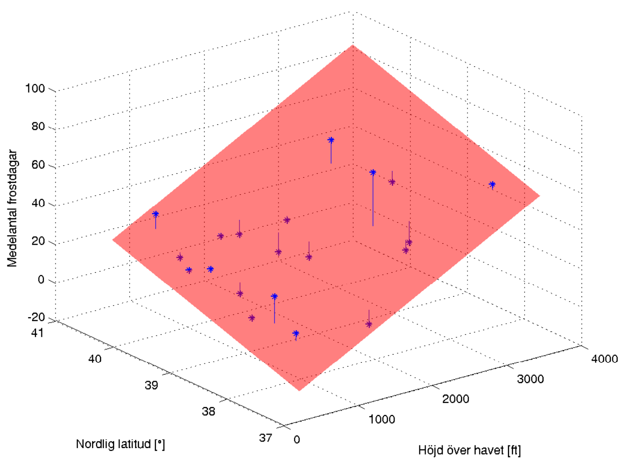
$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{pi} + \varepsilon_i, \quad \varepsilon_i \in \mathrm{N}\left(0, \sigma^2\right) \text{ oberoende}$$

kan skrivas på matrisform som

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \boldsymbol{E}$$

där $\boldsymbol{Y}$ och $\boldsymbol{E}$ är $n \times 1$-vektorer, β en $(p+1) \times 1$-vektor och $\boldsymbol{X}$ en $n \times (p+1)$-matris

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{21} & \cdots & x_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{E} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$



Skattning av β och σ^2 (Kap. 11.3)

MK-skattningar av $\beta_0, \ldots, \beta_p$ (elementen i β) blir

$$\beta^* = (X^\top X)^{-1} X^\top Y \qquad V(\beta^*) = \sigma^2 (X^\top X)^{-1}$$

och skattning av σ^2 är

$$s^2 = \frac{Q_0}{n-(p+1)}$$

där **residualkvadratsumman** ges av

$$Q_0 = \sum_{i=1}^{n} (y_i - \beta_0^* - \beta_1^* x_{1i} - \ldots - \beta_p^* x_{pi})^2$$

Konfidensintervall och hypotestest för β_i

Konfidensintervall för β_i blir alltså

$$I_{\beta_i} = \beta_i^* \pm t_{a/2}(f) \cdot d(\beta_i^*) = \\ = \beta_i^* \pm t_{a/2}(n-p-1) \cdot s\sqrt{[(X^\top X)^{-1}]_{i,i}}$$

Ett **konfidensintervall** för $\mu^*(x_0)$ blir således

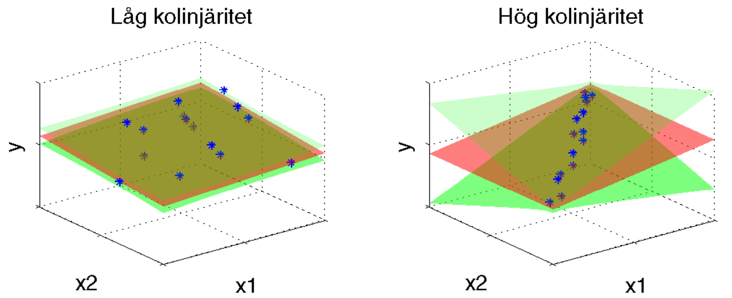
$$I_{\mu^*(x_0)} = x_0\beta^* \pm t_{a/2}(n-p-1)s\sqrt{x_0\left(X^\top X\right)^{-1} x_0^\top}$$

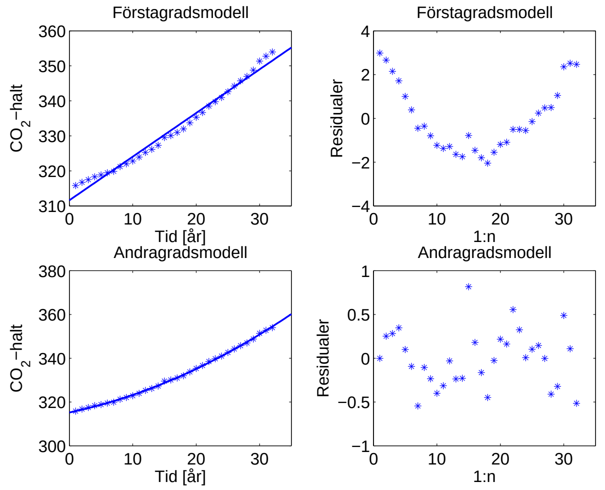
För **prediktionsintervallet** får man, som tidigare, lägga till en etta under kvadratroten

$$I_{Y(x_0)} = x_0\beta^* \pm t_{a/2}(n-p-1)s\sqrt{1 + x_0\left(X^\top X\right)^{-1} x_0^\top}$$

Kolinjäritet (ex. två variabler) (Kap. 11.6)

Man bör om möjligt välja sina (x_{1i}, x_{2i})-värden så att de blir utspridda i (x_1, x_2)-planet och inte klumpar ihop sig längs en linje. Detta ger "en mer stabil grund" åt regressionsplanet.





Linjär $y = \alpha + \beta x$, och kvadratisk, $y = \beta_0 + \beta_1 x + \beta_2 x^2$, anpassning av årlig CO_2-halten vid Mauna Loa som funktion av året (sedan 1960).

Statistiska undersökningar (Kap. 12.1)

Deskriptiv undersökning

Syftar till att beskriva **egenskaper** hos en population.

Analytisk undersökning

Syftar till att undersöka **effekter** av olika **förklarande variabler** eller **faktorer** på en population.

Analytisk undersökning (Kap. 12.1)

Observationsstudie

Ett antal objekt observeras tillsammans med en behandling. Vi har **ingen möjlighet** att påverka behandlingen.

Kontrollerat experiment

Behandlingen av olika objekt kan **kontrolleras** och bestäms på förhand

- Faktorer: Variabler som vi kan **styra** i experimentet
- Kovariater: Variabler som kan **mätas** men inte styras.
- Övriga variabler: Kallas på engelska **confounding factors**

2^2-försök (Kap. 12.3)

I ett 2^2-försök har man 2 **faktorer** som alla kan varieras på 2 **nivåer**.

B	Hög	μ_{12}	μ_{22}
	Låg	μ_{11}	μ_{21}
		Låg	Hög
		A	

Enkel effekt Effekten av **en faktor** om den andra faktorn är **fix**.

Huvudeffekt Effekten av **en faktor** för **alla värden** på den andra faktorn.

Samspelseffekten Skillnaden mellan de **enkla effekterna**.

Teckenschema för 2^2-försök (Kap. 12.3.1)

	Försök	Respons	μ	A	B	AB
A och B låg	(1)	μ_{11}	+	-	-	+
A hög	a	μ_{21}	+	+	-	-
B hög	b	μ_{12}	+	-	+	-
A och B hög	ab	μ_{22}	+	+	+	+

Model för 2^2-försök (Kap. 12.3.2)

Vid n mätningar (**replikat**) av varje faktorkombination ges varje observation av

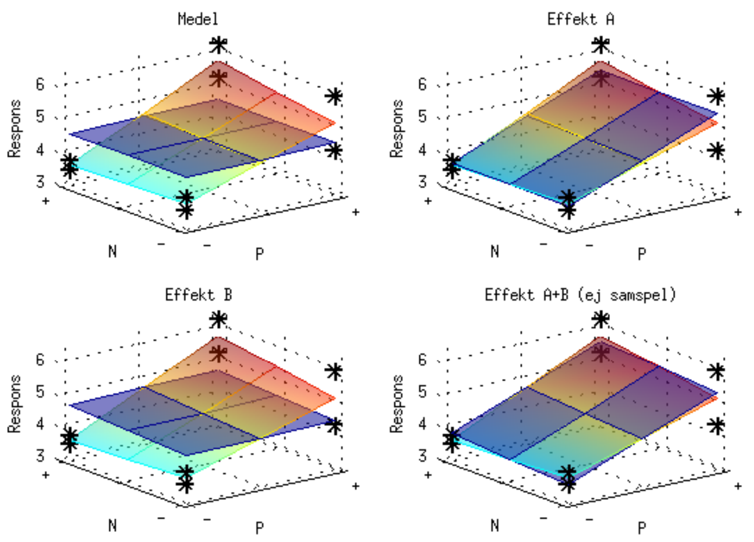
$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk}, \quad i = 1,2;\ j = 1,2;\ k = 1,\ldots,n$$

och felen antas vara **oberoende** $\varepsilon_{ijk} \in \mathrm{N}\left(0, \sigma^2\right)$.

$$\mu_{ij}^* = \bar{y}_{ij\cdot}, \qquad s_{ij}^2 = \frac{\sum_{k=1}^{n} \left(y_{ijk} - \bar{y}_{ij\cdot}\right)^2}{n},$$

$$\widehat{A} = \frac{-\bar{y}_{11\cdot} + \bar{y}_{21\cdot} - \bar{y}_{12\cdot} + \bar{y}_{22\cdot}}{2^2},$$

$$s^2 = \frac{s_{11}^2 + s_{21}^2 + s_{12}^2 + s_{22}^2}{2^2}, \qquad f = 2^2 \cdot (n-1)$$



Konfidensintervall för 2^2-försök (Kap. 12.3.2)

Givet n replikat blir konfidensintervallen för effekterna

$$I_A = \widehat{A} \pm t_{\alpha/2}(2^2 \cdot (n-1)) \frac{s}{\sqrt{2^2 \cdot n}}$$
$$I_B = \widehat{B} \pm t_{\alpha/2}(2^2 \cdot (n-1)) \frac{s}{\sqrt{2^2 \cdot n}}$$
$$I_A B = \widehat{AB} \pm t_{\alpha/2}(2^2 \cdot (n-1)) \frac{s}{\sqrt{2^2 \cdot n}}$$