

HYPOTETESTETING

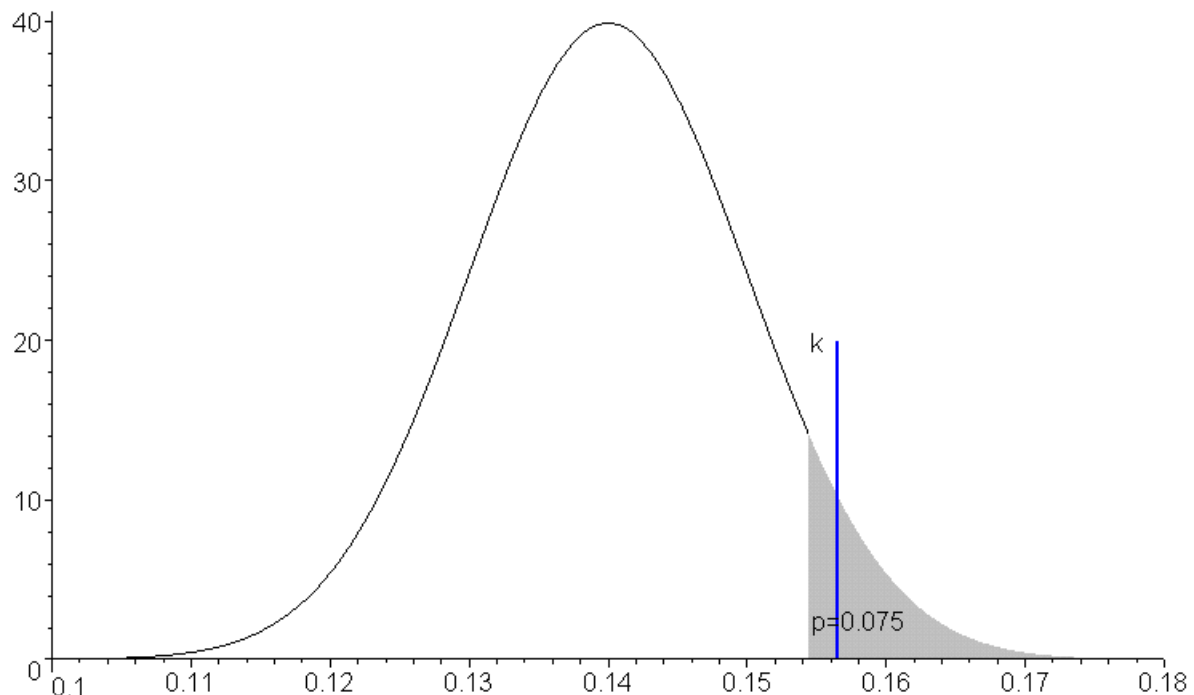
Løsningsforslag fra Tor Slind og Hans Petter Hornæs.

6.20. Her er et vanlig forbrukerproblem: for produsenten er fett og vann billig, kjøtt er dyrt. σ er kjent, gj.snittet $\bar{X}$ er vår estimator for fettprosenten μ . Nullhypotesen er maks.verdien 14 %, alternativet er at fett-prosenten er høyere (tvilen kommer produsenten til gode):

$$\begin{aligned} H_0: \quad \mu &= \mu_0 = 0.14 & \alpha &= 0.05 \\ H_1: \quad \mu &= \mu_1 > 0.14 \end{aligned}$$

Måleverdiene antas å følge H_0 's normalfordeling $N(0.14, 0.03)$. Gjennomsnittet $\bar{X}$ av 9 målinger er normalfordelt $N(0.14, 0.01)$. Forkaster H_0 hvis gjennomsnittet er **større enn** (se H_1) en kritisk verdi k beregnet fra normalfordelingens kvantil $z_\alpha = z_{0.05} = 1.645$, se neste figur:

$$k = \mu_0 + z_\alpha \sigma_{\bar{X}} \Rightarrow k = 0.14 + 1.645 \frac{0.03}{\sqrt{9}} = \underline{0.15645}$$



Gjennomsnittsverdien:

$$\bar{x} = \frac{0.14 + 0.15 + 0.18 + 0.12 + 0.17 + 0.13 + 0.15 + 0.19 + 0.16}{9} = 0.15444$$

Konklusjon: $\bar{x} < k$, kjøttdeigen inneholder ikke for mye fett.

p -verdien er sannsynligheten for å få en gj.snittsverdi større enn den vi har observert, bruker $N(0.14, 0.01)$:

$$p = P(\bar{X} > \bar{x}) = P\left(\frac{\bar{X} - \mu}{\sigma_{\bar{X}}} > \frac{0.1544 - 0.14}{0.01}\right) = 1 - \Phi(1.44) = \underline{0.075}$$

Dersom $p > \alpha$, er måleverdien $\bar{x}$ nedenfor k , og vi kan ikke forkaste H_0 , se forrige figur.

Styrkefunksjonen er sannsynligheten for å velge $H1$ ($\bar{x} > k$) hvis det virkelig er for mye fett i deigen, altså sanns. for å velge riktig konklusjon hvis $H1$ gjelder. Dette uttrykkes som

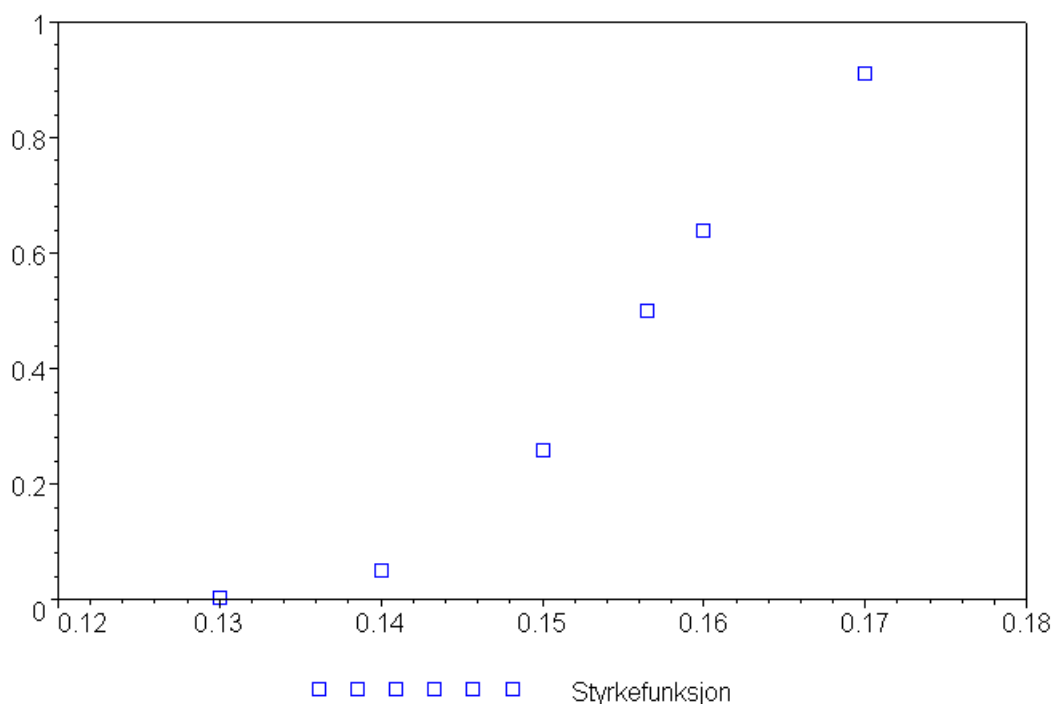
$$\gamma(\mu_1) = P(\bar{X} > k \mid \mu = \mu_1) = P\left(\frac{\bar{X} - \mu_1}{\sigma_{\bar{X}}} > \frac{0.15645 - \mu_1}{0.01}\right) = 1 - \Phi\left(\frac{0.15645 - \mu_1}{0.01}\right)$$

Setter inn forskjellige μ -verdier og tegner opp hvordan styrkefunksjonen endrer seg. Uten å regne på nytt kjenner vi allerede 2 verdier: $\gamma(\mu_0) = \alpha = 0.05$ og $\gamma(\mu_1 = k) = 0.5$.

μ_1	0.13	0.14	0.15	$k=0.15645$	0.16	0.17
$\gamma(\mu_1)$	0.004	0.05	0.26	0.50	0.64	0.91

Med 9 målinger er sanns. for å velge feil hypotese (velge $H0$) når $\mu = 0.16 > k$ lik $1 - 0.64 = 0.36$, se tabellen ovenfor. Hvis vi øker antall målinger vil k avta, styrkefunksjonen får en større verdi for $\mu = 0.16$, og sanns. for et feil valg blir redusert. Vi kan innføre antall målinger i uttrykket for styrkefunksjonen, som skal ha verdien $\gamma(0.16) = 1 - 0.10 = 0.9$:

$$\begin{aligned} \gamma(0.16) &= 1 - \Phi\left(\frac{k - 0.16}{\sigma / \sqrt{n}}\right) = 1 - \Phi\left(\frac{0.14 + 1.645\sigma / \sqrt{n} - 0.16}{\sigma / \sqrt{n}}\right) = 0.9 \\ \Phi\left(\frac{0.14 + 1.645\sigma / \sqrt{n} - 0.16}{\sigma / \sqrt{n}}\right) &= 0.1 \Rightarrow \frac{0.14 + 1.645\sigma / \sqrt{n} - 0.16}{\sigma / \sqrt{n}} = -1.282 \\ -0.02\sqrt{n} &= \sigma(-1.282 - 1.645) = -0.08781 \Rightarrow n > 19.27 = \underline{20} \end{aligned}$$



6.21. Må bruke $\bar{x}$ og s fra 9 målinger:

$$\bar{x} = 103.78, \quad s = 2.949$$

$H0$ er maks.grensen, alternativet om støynivået er større enn 100 dB. Bruker et 5 % signifikansnivå på testen:

$$\begin{aligned} H_0: \quad \mu &= \mu_0 = 100 & \alpha &= 0.05 \\ H_1: \quad \mu &= \mu_1 > 100 \end{aligned}$$

Tester om $\bar{x}$ er større en grenseverdien $k = \mu_0 + t_{0.05,8}s/\sqrt{9}$ ($t_{0.05,8}$ er # std.avvik fra t-fordelingen):

$$k = 100 + 1.86 \frac{2.949}{3} = 101.8 \Rightarrow \bar{x} = 103.8 > k$$

Vi forkaster H_0 , støynivået var for høyt.

Eventuelt:

I) Observert gj.snitt gir t-verdien

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{3.78}{0.983} = 3.845 > t_{0.05,8} \Rightarrow \text{Forkaster } H_0$$

II) Observert p -verdi:

$$p = P(\bar{X} > \bar{x}) = P\left(T > \frac{\bar{x} - \mu_0}{s/\sqrt{n}}\right) = P(T > 3.845) = 0.00246 < \alpha \Rightarrow \text{Forkaster } H_0$$

6.22. Starter med konfidensintervallet, må bruke t-fordelingen siden σ er ukjent:

$$\bar{x} = 110\,365 \quad s = 1044$$

Et 95 % konf.intervall er gitt av 2.5%-kvanteren $t_{0.025,9} = 2.262$ (antall standardavvik):

$$\left[110\,365 - 2.262 \cdot \frac{1044}{\sqrt{10}}, 110\,365 + 2.262 \cdot \frac{1044}{\sqrt{10}} \right] = [109\,618, 111\,112] [\text{tonn}]$$

Dette resultatet kan vi sammenligne med en hypotesetest hvis vi bruker et signifikansnivået $\alpha = 2.5$ %. Utregning lik oppg. 19:

$$\begin{aligned} H_0: \quad \mu &= \mu_0 = 110\,000 & \alpha &= 0.025 \\ H_1: \quad \mu &= \mu_1 > 110\,000 \end{aligned}$$

Tester om $\bar{x}$ er større en grenseverdien $k = \mu_0 + t_{0.025,9}s/\sqrt{10}$:

$$k = 110\,000 + 2.262 \frac{1044}{\sqrt{10}} = 110\,747 \Rightarrow \bar{x} = 110\,365 < k$$

Konklusjoner:

- 1) Konfidensintervallet inneholder verdien 110 000. Det er 95 % sannsynlig at forventningsverdien ligger i intervallet [109 600, 111 100] [tonn].
- 2) Siden $\bar{x} < k$ kan vi ikke forkaste H_0 ; ingen grunn til å tro at forventningsverdien er større enn 110 000 tonn.

6.29. Estimerer defektandelen p med estimatoren $\hat{p} = \frac{X}{n}$, X er antall defekte enheter. X er binomisk fordelt, $X \sim \text{bin}(134, p)$. $\hat{p}$ er forventningsrett, se oppg. 6.5 og 6.14:

$$E(\hat{p}) = p, \quad \text{Var}(\hat{p}) = \frac{p(1-p)}{n}$$

Erstatter p med $\hat{p}$ i variansen for X og sjekker om vi kan bruke $N(p, (\hat{p}(1-\hat{p})/n)^{1/2})$ som en tilnærming for den binomiske fordelingen:

$$Var(X) \approx n\hat{p}(1-\hat{p}) = 134 \frac{6}{134} \frac{128}{134} = 5.731 > 5$$

Siden variansen til X er litt større enn 5, er normalfordelingen en brukbar tilnærming:

$$\hat{p} \sim N(p, 0.01787)$$

Hypotesetesten blir:

$$H_0: p = p_0 = 0.1 \quad \alpha = 0.05$$

$$H_1: p = p_1 \neq 0.1$$

Kommentar: H_1 velges ut fra at hun verken må lage for få eller for mange vrak.

Forkastningsgrensene k_1 og k_2 er bestemt ved (σ antas kjent)

$$k_{1,2} = 0.10 \pm z_{0.05/2} \sigma_{\hat{p}} = 0.10 \pm 1.96 \cdot 0.01787 = \begin{cases} 0.065 \\ 0.135 \end{cases}$$

Observert defektprosent er: $\hat{p}_{obs} = \frac{6}{134} = 0.045$

Konklusjon: Den observerte feilandelen 0.045 er mindre enn laveste grense 0.065, vi velger H_1 . Gunnhild produserer for få vrak.

Alternative måter til å finne testresultatet er vist i oppg. 19. I denne oppgaven må vi bytte ut testobservatoren T med Z siden vi antar at standardavviket er kjent.