Statistikk

Jo Eidsvik

Matematiske fag, NTNU

Kontinuerlige fordelinger

Stokastisk variabel X. Sannsynlighets**tetthet** $f(x)$.

$$\int_{x \in S} f(x)dx = 1, \qquad f(x) \geq 0.$$

Normalfordeling

Den mest brukte fordelingen i statistikk.
Tetthet:

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

- Symmetrisk om $\mu = E(X)$. Variasjonen styres av $\sigma^2 = Var(X)$.
- Notasjon: $X \sim N(\mu, \sigma^2)$.

Normalfordeling

- $X \sim N(\mu, \sigma^2)$. Parametre er $\mu = E(X)$ og $\sigma^2 = Var(X)$.
- Fine skift- og skaleringsegenskaper: $Z = (X - \mu)/\sigma$, $Z \sim N(0, 1)$

$$f(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right).$$

  Ca 95 prosent av sannsynlighet inne (-2,2). Ca 68 prosent innen (-1,1).
- Approksimerer mange andre fordelinger (når datastørrelse, n, blir stor).

Dear Abby, 266 days

Dear Abby: You wrote in your column that a woman is pregnant for 266 days. Who said so? I carried my baby for ten months and five days, and there is no doubt about it because I know the exact date my baby was conceived. My husband is in the Navy and it couldn't have possibly been conceived any other time because I saw him only once for an hour, and I didn't see him again until the day before the baby was born.

I don't drink or run around, and there is no way this baby isn't his, so please print a retraction about the 266-day carrying time because otherwise I am in a lot of trouble.

San Diego Reader

Dear Abby, 266 days

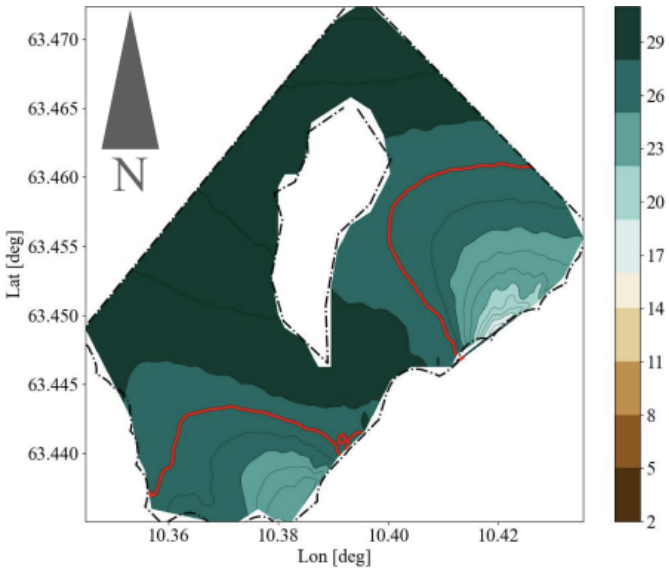
Lengde på svangerskap:

- $\mu = 266$, $\sigma = 16$, Normalfordelt
- $P(X > 310) = 1 - P(X < 310) = 1 - P(Z < \frac{310-266}{16}) = 1 - P(Z < 2.75) = 0.003$

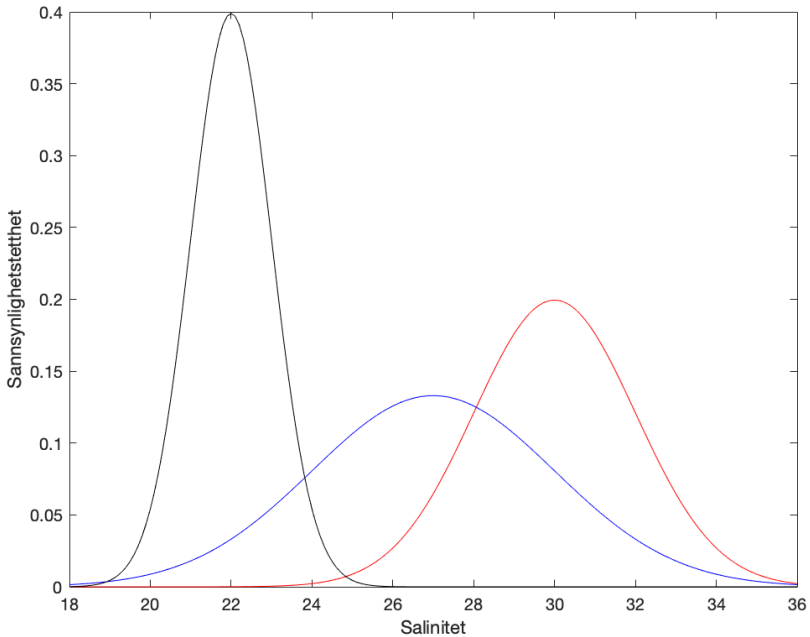
310 dager eller mer er veldig lenge.

Salt i fjorden





Salt i fjorden



Antar salt ved bøye utenfor Munkholmen $X \sim N(30, 2^2)$.

$$P(X > 35) = P\left(\frac{X - 30}{2} > \frac{35 - 30}{2}\right) = P(Z > 2.5) = 0.006$$

IOT anvendelse : SMS sendes fra bøye



Send SMS om måling er veldig lav. I lange løp - kun 1 av 50 dager:

$$P(X < t) = 0.02$$

Finner threshold t for å sende måling:

$$P\left(\frac{X-30}{2} < \frac{t-30}{2}\right) = 0.02$$

$\frac{t-30}{2} = -2.05$
$t = -2.05 \cdot 2 + 30 = 25.9$