# Repetisjon av P-verdier og hypotesetester for forventningsverdi og andeler

- Anta at vi har
  - Populasjon med forventningsverdi $\mu$ og varians $\sigma^2$.
  - Tilfeldig utvalg: $X_1, X_2, \ldots, X_n$.
  - $n \geq 30$ eller normalpopulasjon.

  Hvis vi ønsker å teste

  $$H_0 : \mu = \mu_0$$
  $$H_1 : \mu \neq \mu_0$$

  ved signifikansnivå $\alpha$ bruker vi forkastingsregelen

  - Hvis $\sigma$ er **kjent**: forkast hvis $z = (\bar{x} - \mu_0)/(\sigma/\sqrt{n}) > z_{\alpha/2}$ eller $z < -z_{\alpha/2}$.
  - Hvis $\sigma$ er **ukjent** og **normalpopulasjon**: forkast hvis $t = (\bar{x} - \mu_0)/(s/\sqrt{n}) < t_{\alpha/2,n-1}$ eller $t < -t_{\alpha/2,n-1}$.

- **Ensidige hypotesetester** $H_1 : \mu < \mu_0$ eller $H_1 : \mu > \mu_0$ oppnås ved å velge $\alpha$ sannsynlighet i nedre eller øvre hale. For eksempel for $H_1 : \mu < \mu_0$ med kjent $\sigma$ forkaster man hvis $z < -z_\alpha$.

- **MERK**: Det at hypotesetesten har signifikansnivå $\alpha$ betyr at sannsynligheten for en Type I feil er $\alpha$.

- En **P-verdi** er den høyeste verdien av signifikansnivået der den observerte verdien av testobservatoren er signifikant.

- En **P-verdi** kvantifiserer styrken av bevisene mot nullhypotesen. Lavere verdi betyr sterkere bevis mot nullhypotesen. Fordelen er at i steden for bare ja/nei-konklusjonen har man informasjon om styrken av bevis som lå til grunn for konklusjonen.

- En vanlig hypotesetest svarer til å regne ut P-verdi og forkaste hvis P-verdi er mindre enn $\alpha$.

- Anta at vi har
  - Populasjon der en ukjent andel $p$ har en bestemt egenskap.
  - Tilfeldig utvalg: $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Bernoulli}(p)$.

  Vi ønsker en ensidig eller tosidig test på denne andelen med $H_0 : p = p_0$. Da er den relevante testobservatoren $X \sim \text{Binomial}(n, p)$, og man har to muligheter for å utføre testen:

  - å bruke eksakt fordeling når $n$ er liten. Her regner man ut P-verdi og sammenligner med ønsket signifikansnivå $\alpha$.
  - å bruke en tilnærming med normalfordeling når $n$ er stor. Bruk en $z$-test basert på

    $$Z = \frac{X - np_0}{\sqrt{np_0(1 - p_0)}} \sim \mathcal{N}(0, 1),$$

    med vanlige forkastingsregler

- **VIKTIG:** Det er ikke mulig å bevise at $H_0$ er usann. Å forkaste $H_0$ med signifikansnivå $\alpha$ er et utsagn om at bevisene mot $H_0$ er så sterke at man er villig til å konkludere med at $H_0$ er usann. Hvor sterke bevisene må være, det vil si hvor liten $\alpha$ må være, kan avhenge både av personen som utfører analysen og av de potensielle konsekvensene av å trekke feil konklusjon.