# SF1914/SF1916: Sannolikhetsteori och statistik
# Föreläsning 12
# Hypotesprövning.

Tatjana Pavlenko

1 oktober 2018

KTH Matematik

# Plan för dagens föreläsning

- Intervallskattning med normalfördelade data: två stickprov (rep.)
- Intervallskattning med normalapproximation (Kap. 12.4)
- Begrepp inom hyptesprövning (Kap. 13.2-13.3).
- Testvariabel- och direktmetoden.
- Tillämpning på normalfördelning (Kap. 13.6).

## INTERVALLSKATTNING MED NORMALFÖRDELADE DATA: TVÅ STICKPROV (REP.)

Under förra föreläsning fick vi konfidensintervall för två typ av situationer med två stickprov.

- *Två oberoende stickprov*: $I_{\mu_1-\mu_2}$ dvs intervall för differensen mellan två väntevärdena hos två normalfördelningar $N(\mu_1, \sigma_1)$ och $N(\mu_2, \sigma_2)$ i fallen då
    1. $\sigma_1$ och $\sigma_2$ var kända
    2. $\sigma_1 = \sigma_2 = \sigma$ och okänd.
- *Stickprov i par*: $I_\Delta$, dvs intervall för systematisk skillnad $\Delta$ baserat på differenser $z_i = x_i - y_i$, där $x_i$ och $y_i$ är oberoende observationer från $N(\mu, \sigma_1)$ respektive $N(\mu + \Delta, \sigma_2)$. Här kan standardavvikelserna $\sigma_1$ och $\sigma_2$ vara olika.

- T ex för fall 2 får vi:

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \in t(n_1 + n_2 - 2),$$

Från formelsamling: 11.2 (a)+11.2 (d)+12.2 ($t$-metoden)

$$I_{\mu_1 - \mu_2} = (\bar{x} - \bar{y} \pm t_{/2}(k)d),$$

med $d = s\sqrt{1/n_1 + 1/n_2}$ och $k = n_1 + n_2 - 2$.

# Stickprov i par (rep.)

- Vi studerar $n$ st objekt och på varje objekt gör mättningar *före* och *efter* ... Detta ger data
  - $x_1, \ldots, x_n$ som är obesrvationer av $X_i \in N(\mu_i, \sigma_1)$, ober. och
  - $y_1, \ldots, y_n$ som är obesrvationer av $Y_i \in N(\mu_i + \Delta, \sigma_2)$, ober., $i = 1, \ldots, n$.
  - Vi tolkar $\Delta$ som förändringen.
- Vi bygger $z_i = y_i - x_i$ som betraktas som observationer av $Z_i = Y_i - X_i \in N(\Delta, \sigma_z$ med $\sigma_z = \sqrt{\sigma_1^2 + \sigma_2^2}$. Vidare
  - $\bar{Z} \in N(\Delta, \frac{\sigma_z}{\sqrt{n}}$ om $\sigma_z$ är känd, (formls. 11.1 (a))
  - $\frac{\bar{Z} - \Delta}{S_z/\sqrt{n}} \in t(n-1)$ om $\sigma_z$ är okänd, (formls. 11.1 (d))
- Från formelsamling, med 11.1 (a)+11.1 (d)+12.2 ($t$-metoden) får vi ett konfidensintervall med konf. grad $1 - \alpha$

$$I_\Delta = (\bar{z} \pm t_{\alpha/2}(n-1)d),$$

där $d = s_z/\sqrt{n}$.

# Intervallskattning med normalapproximation.

- Idé: Vi har hittils studerat olika situationer där obrservationerna, och därmed även skattarre, är *exakt* normalfördelade.
- Men även om vi inte har stickprov från en normalfördelning, det mycket väll på grund av Centrala gränsvärdessatsen, kan vara så att vår skattare är *approximativt* normalfördelad.
- Vi kan i så fall tillämpa normalfördelningsteori, men istället för att få exakt konfidensgrad $1 - \alpha$ så får vi nöja oss med att att konfidensgrad blir *approximativt* $1 - \alpha$.

KTH Matematik

## Intervallskattning med normalapproximation (fort.).

- Låt $x_1, \ldots, x_n$ vara ett stickprov från någon s.v. $X$ med någon fördelning som beror på okänd parameter $\theta$.
- Låt $\theta^*$ vara en punktskattare av $\theta$ sådan att

$$\theta^* \in AsN(\theta, D).$$

- Om $D$ är känd så får vi, på analogt sätt med $\lambda$-metoden konfidensintervall för $\theta$

$$I_\theta = (\theta^*_{obs} \pm \lambda_{\alpha/2} D).$$

- Oftast $D$ är okänd och vi skattar den som tidigare med medelfelet $d$, villket innebär en ytterliggare approximation, men intervallet

$$I_\theta = (\theta^*_{obs} \pm \lambda_{\alpha/2} d)$$

är fortfarande ett konfidensintervall med *approximativ konfidensgrad* $1 - \alpha$.

KTH Matematik

# Intervallskattning med normalapproximation (fort.).

- Exempel (på tavlan): En tillverkare säljer komponenter i förpackningar med 50 komponenter i varje. Komponenterna är defekta med sannolikhet $p$ oberoende av varandra. En kund köper en förpackning och konstaterar att 15 komponenter var defekta ( de övriga 35 var hela). Gör ett tvåsidigt 95% approximativt konfidensintervall för $p$.

- Svar: $(0.17, 0.43)$.

KTH Matematik

# Hypotesprövning

- Då man konstruerar konfidenintervall vill man dra slutsatser om okända parametrar från data. Detta kan också kallas för *singifikans test* eller *hypotesprövning*.
- Vi börjar med
- Exempel: *Skicklig slantslingning.*

  En person (se bild på nästa sida) påstår sig vara tillräckligt fingerfärdig för att kunna påverka vilken sida som kommet upp vid slantslingning, så att han får krona oftare är klave, dvs med $p = P(krona) > 1/2$.

- Hur skulle kunna ett statsitiskt experiment och analys gå till?
- Vi är skeptiska mot hans påstående och vi är beredda att tro honom (dvs *förkasta hypotesen att* $p = 1/2$ *om han får tillräckligt många krona.*

## SKICKLIG SLANTSLINGNING?


FIGUR: Personen som inspirerade exempel ovan en amerikanske statistik professorn Persi Diaconis, som före statistikkarriären livnärde sig som magiker. Han har bl. a. gjort ingående studier av slantslingning.

Läs mer om Persi Diaconis här

- https://www.quantamagazine.org/20150414-for-persi-diaconis-next-magic-trick/
- https: PersiDiaconis.pdf

## Skicklig slantslingning? (forts.)

- För att testa påståendet får personen göra 10 kast. Låt $x$ vara antalet krona vid försöket. Vårt test bör då vara av typen:

  *Förkasta hypotesen att $p = 1/2$ om $x \geq C$*

  för ett lämpligt vall av konstanten $C$. Vi vill välja $C$ så att risken att vi av misstag förkastar hypotesen $p = 1/2$ blir liten.

- Vi noterar att $x$ är en observation av en s.v. $X \in Bin(10, p)$. Vi behöver alltå studera $P(X \geq C)$ då $p = 1/2$ för olika värden på konstanten $C$.

# Skicklig slantslingning? (forts.)

- $C = 10$ ger

$$P(X \geq C) = \binom{10}{10}\left(\frac{1}{2}\right)^{10} = \frac{1}{1024} \approx 0.000098.$$

- $C = 9$ ger

$$P(X \geq C) = \binom{10}{10}\left(\frac{1}{2}\right)^{10} + \binom{10}{9}\left(\frac{1}{2}\right)^{10} = \frac{11}{1024} \approx 0.011.$$

- $C = 8$ ger

$$P(X \geq C) = \left(\binom{10}{10} + \binom{10}{9} + \binom{10}{8}\right)\left(\frac{1}{2}\right)^{10} = \frac{56}{1024} \approx 0.055.$$

- Om vi vill hålla felrisken liten (t ex. $\leq 0.05$) verkar $C = 9$ vara ett rimligt val. Beslutsregel blir då att

  *förkasta hypotesen att $p = 1/2$ om antalet krona är minst nio,*

  och felrisken blir 1.1%.

KTH Matematik

## BEGREPP INOM HYPOTESPRÖVNING.

- Vi betraktar den allmänna situationen: vi har ett stickprov $x = (x_1, \ldots, x_n)$ som är utfall av s. v. $(X_1, \ldots, X_n)$ från någon fördelning $F_X(x; \theta)$.
- Vi vill testa en grundhypotes, eller *nollhypotes*, om $\theta$: $H_0 : \theta = \theta_0$ mot en *alternativ hypotes*, $H_1$, som kan vara *enkel*, $H_1 : \theta = \theta_1$, eller *sammansatt*, t ex $H_1 : \theta > \theta_0$. En hypotes av typ $H_1 : \theta > \theta_0$ eller $H_1 : \theta < \theta_0$ är *ensidig*, medan $H_1 : \theta \neq \theta_0$ är *tvåsidig*.
- För att testa $H_0$ definierar vi först en *test variabel* eller *teststorhet* , $t_{obs} = t(x)$ som är en observation av motsvarande stickprovsvariabel $t(X)$. Vi kommer också att behöva testets signifikansnivån eller felrisk.
- Def: *Med signifikansnivån*, eller *felrisken,* $\alpha$ för ett test menas

$$\alpha = P(H_0 \text{ förkastas}) \quad \text{om } H_0 \text{ är sann.}$$

- Testet utformas på så sätt att felrisken blit liten. Vanliga nivåer är $\alpha = 0.05$, $\alpha = 0.01$, $\alpha = 0.001$. Jämför med felrisker vid konfidensintervall konstruktion!

## Testvariabelmetoden.

- Hitta en teststorhet $t(X)$ och till den *ett kritiskt område* $C$.
- Testet blir då:

$$\text{Om} \begin{cases} t_{obs.} \in C, & \text{förkasta } H_0 \\ t_{obs.} \notin C, & \text{ej förkasta } H_0, \end{cases}$$

  där det kritiska området $C$ anpassas så att

$$P(H_0 \text{ förkastas}) = P(t(X) \in C) = \alpha \quad H_0 \text{ är sann}$$

  och $\alpha$ är testets signifikansnivå som väljs på förhans.
- Tolkning:
  *Om $H_0$ förkastas så föreligger det ett signifikant avvikelse från nollhypotesen $H_0$ på nivån $\alpha$.*

KTH Matematik

- Låt i detta fall $x = 7$ (dvs personen fick krona 7 ggr ut av 10 kast), $x$ är en observation av $X \in Bin(10, p)$.
- Vi låter nollhypotesen vara

$$H_0 : p = p_0 = 1/2,$$

dvs vi är skeptiska mot pesonens påstående att kunna påverka vilken sida som kommer upp vid slantsingling och tror att $p = 1/2$. Mothypotesen är

$$H_1 : p > p_0,$$

dvs personen kan påverka resultatet.

- Testet utförs på signifikansnivån $\alpha = 0.01$.

# Exempel: Skicklig slantslingning? (forts.)

- Vi har i detta fall en enda observation så att $t_{obs} = t(x) = x$.
- $C$ ska väljas så att följande uppfylls:

  $$P(H_0 \text{ förkastas}) = P(T(X) \in C) = P(X \in C) = \alpha, \text{ om } H_0 \text{ är sann},$$

  och vi väljer $C = [x_\alpha, 10]$ där $x_\alpha$ är 0.01-kvantilen för $Bin(10, 0.5)$-fördelningen. Om ett utfall $x$ hamnar i kritisk område $C$ så förkastas $H_0$ till förmån för $H_1$, dvs vi tror på att $p$ är större än $1/2$.
- Med hjälp av MATLAB får man $x_{0.01}$-kvantilen enligt

```
>> binoinv(0.99,10,0.5)
ans =
9
```

- Slutsats: Då personen fick 7, och $x = 7 \notin C = [9, 10]$ så kan vi ej förkasta $H_0$ att $p = 1/2$.

## Exempel: test av astrolog.

- Detta exempel är baserat på ett verkligt fall. I Lund framlades en psykologuppsats (motsv exjobb) som påstāds att *horoskop ger information om personligheten.* En astrolog påstår att han kan utläsa en persons framtid från utseendet på stjärnhimlen då personen föds.
- Astronomen från Chalmers, Curt Roslund organiserade ett experiment där 15 dömda mördare och 15 sekreterare väljs ut och astrologen får tillgång till bilder av stjärnhimlen då de 30 personerna föddes. Uppgiften bestād i att peka ut så många mördare som möjligt.
- Hur skulle en statistisk analys av experimentet gå till? Syftet är att konstruera ett test för att pröva hypotesen

$$
\begin{aligned}
H_0 : &\quad \text{astrologen bara gissar} \\
H_1 : &\quad \text{det inte är så att han bara gissar.}
\end{aligned}
$$

- Vid experimentet lyckades astrologen utpeka rätt 9 mördare. Kan $H_0$ förkastas? förkastas? På tavla i mån av tid med *Direktmetoden*, se nästa slide. (Svar: $H_0$ kan ej förkastas).

# Direktmetoden.

- Ett alternativ till testvariabel metoden.
- Metodens idé med hjälp av exempel *Skicklig slantslingning* (på tavlan).
- Allmänt: Det *direkta* med direktmetoden är att vi räknar ut ett storhet som direkt kan jämföras med felrisken! Metoden går till på följande sätt:

  1. Antag att $H_0$ är sann (dvs att det parametervärde den specifierar är det rätta värde).
  2. Räkna under $H_0$ ut
  $$p\text{-värdet} = P(\text{teststorhet blir minst lika extremt som observerat}).$$
  3. Om $p$-värde är mindre än $\alpha$ (ofta $\alpha = 0.05, 0.01, 0.001$) så förkastas $H_0$ med felrisk $\alpha$.

- Tolkning: Om $p$-värde är litet (ofta $<$ än $0.05, 0.01$ eller $0.001$) så tror vi inte på $H_0$. $p$-värde ger alltså ett mått på *hur orimlig $H_0$ är.*

KTH Matematik

# Nästa föreläsning.

- Konfidensmetoden: hur kan man utföra test med konfidensintervall?
- Tillämpning på normalfördelning: två stickprov.
- Styrkefunktion
- Normalapproximation.

KTH Matematik