8 Inferens om väntevärdet (och variansen) av en fördelning

8.2 Skattning av μ och Students T-fördelning

Om σ är känd, kan man använda statistikan

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

för att hitta konfidensintervall för μ. Om σ inte är känd, måste man skatta den.

1) Hur kan man skatta σ? Genom att använda stickprovsstandardavvilkelsen

$$\hat{\sigma} = S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$$

2) Vad blir fördelningen av den ovanstående statistikan om man ersätter σ med S? Statistikan är inte längre normalfördelad men har en så-kallad **Students T-fördelning**.

Egenskaper av T-fördelningen

1) Fördelningen har en parameter γ, som är ett positivt heltal och som kallas **antal frihetsgrader** (degree of freedom). Man betecknar T-fördelningen med γ frihetsgrader av T_γ.

2) Den är en kontinuerlig fördelning som är "klockformad" som normalfördelningen. Väntevärdet av T-fördelningen är 0.

3) Parametern γ påverkar formen (variansen) av fördelningen

   - när γ blir större, blir variansen mindre
   - när $\gamma \to \infty$, närmar T-fördelningen sig standardiserade normalfördelningen

Sats 8.2.1: Låt $X_1, X_2, ..., X_n$ vara ett stickprov från normalfördelningen med väntevärde μ och varians σ^2. Då är

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim T_{n-1}.$$

Statistikan ovan kan man använda för att hitta konfidensintervall för μ då variansen σ^2 inte är känd. $100(1-\alpha)\%$ konfidensintervall för μ är

$$\bar{X} \pm t_{\frac{\alpha}{2}} S/\sqrt{n},$$

där $t_{\frac{\alpha}{2}}$ är punkten associerad till T_{n-1}-fördelningens täthetskurva så att ytan under kurvan till höger av punkten $t_{\frac{\alpha}{2}}$ är $\frac{\alpha}{2}$ (dvs. att $P(T_{n-1} > t_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$).

Ex. 8.2.2: Svaveldioxid (sulfur dioxide) och kväveoxid (nitrogen oxide) får man när man använder fossilbränsle. Dessa föreningar kan resa långa avstånd och de kan förvandlas till syror innan de kommer ner som syrt regn. Nu har man 24 mätningar av SO_2-konsentrationen (enhet $\mu g/m^3$) från en skog i Bovaria, Tyskland, som har blivit skadad av syrt regn. Man vill hitta 95% konfidensintervall för medel-SO_2-konsentrationen i denna skog.

Data:

52.7	43.9	41.7	71.5	47.6	55.1
62.2	56.5	33.4	61.8	54.3	50.0
45.3	63.4	53.9	65.5	66.6	70.0
52.4	38.6	46.1	44.4	60.7	56.4

8.3 Hypotestest

Vad är hypotestest för någonting?

Man observerar någonting och vill generalisera det man ser/observerar. Man pratar om våra observationer generellt

1) Ett nytt läkemedel förlängar livet av leukemipatienter

2) Efter två drinkar är medelnivån av alkohol i blodet över den tillåtna nivån

3) Majoriteten av väljarna tycker att premiärministern är ok

Man generaliserar även om man har bara några observationer. Men

- man borde inte generalisera om man inte har tillräckligt mycket bevis

- statistiskt hypotestest hjälper oss att avskilja "bra" generaliseringar från de "dåliga".

Ex.8.3.1: Det finns många faktorer som påverkar hur bra man ser de reflekterade vägmärken på motorvägar. En viktig sak är att bilens framlyktor borde vara rätt siktade. Man tror att mer än 50% av bilar har felsiktade framlyktor. Om man lyckas att stödja det här (att mer än 50% av bilar har felsiktade framlyktor) statistiskt, kommer man att få ett nytt bättre kontrollsystem. (Låt p vara andelen bilar med felsiktade framlyktor.) Man vill testa om majoriteten av bilar har felsiktade framlyktor. Hur kan man göra det?

Två hypoteser:

1) Hypotesen som forskaren föreslår/tror på: majoriteten av bilar har felsiktade framlyktor. Den kallas en **mothypotes** och betecknas av H_1.

$$H_1 : p > 0.5$$

2) Hypotesen som är motsats till forskarens hypotes: andelen bilar med felsiktade framlyktor är högst 50%. Den kallas en **nollhypotes** och betecknas av H_0.

$$H_0 : p \leq 0.5$$

Man antar att H_0 är sann och den blir sann tills man hittar tillräckligt mycket bevis för att förkasta den.

För att testa om H_0 är sann behöver man en statistika och det är naturligt att använda antalet bilar med felsiktade framlyktor X. Först antar man att H_0 är sann, dvs att $p \leq 0.5$. Sedan kollar man hur mycket andelen felsiktade framlyktor X/n (n är stickprovsstorleken) skiljer sig från 0.5 (nollhypotesvärdet som är närmast till värdena i H_1). Om vi kan förkasta hypotesen att $p = 0.5$, då kan man också förkasta hypotesen $p \leq 0.5$. Om värdet av teststatistikan är stor, betyder det att X är större än 0.5. Hur stor borde värdet vara för att kunna förkasta H_0?

I vårt fall är teststatistikan binomialfördelad med parametrar n (stickprovsstorleken, antalet bilar man undersöker) och succésannolikheten 0.5 om H_0 är sann. Då är $\mathbf{E}[X] = n \cdot 0.5$. Säg att man undersöker 20 bilar ($n = 20$). Då är $\mathbf{E}[X] = 10$ vilket betyder att det genomsnittliga antalet bilar med felsiktade framlyktor bland 20 bilar är 10.

Säg att teststatistikan får ett stort värde och att vi bestämmer oss att förkasta H_0. Då är antingen

1) H_0 sann och man har fått ett sällsynt stickprov

eller

2) H_1 sann och majoriteten av bilar har felsiktade framlyktor

Vanligen tror man då att 2) är sann och förkastar H_0. Men vad om 1) är sann? Man har gjort ett felt beslut: fel av typ I (H_0 förkastas även om den är sann). Sannolikheten av fel av typ I

$$P(\text{fel av typ I}) = P(H_0 \text{ förkastas}|H_0 \text{ sann}) = \alpha$$

kallas **signifikansnivån** av testet.

I vårt test:

$H_0 : p \leq 0.5$ och $H_1 : p > 0.5$

Teststatistika X (antalet bilar med felsiktade framlyktor) har $Bin(20, 0.5)$-fördelning, om H_0 är sann

Man förkastar H_0 om $X \geq 14$ och då är $\alpha \approx 0.05$. Värdena av teststatistikan för vilka H_0 förkastas kallas ett **kritiskt område**.

Säg att värdet av teststatistikan är inte tillräckligt stort för att förkasta H_0 och att man accepterar H_0. Det är möjligt att beslutet är felt. Vad om H_0 inte är sann? Då skulle vi acceptera H_0 även om den inte är sann (fel av typ II). Sannolikheten av fel av typ II är

$$P(\text{fel av typ II}) = P(\text{acceptera } H_0 | H_1 \text{ sann}) = \beta$$

Sannolikheten β används för att definiera **teststyrkan**

$$P(H_0 \text{ förkastas} | H_1 \text{ sann}) = 1 - P(\text{acceptera } H_0 | H_1 \text{ sann}) = 1 - \beta.$$

Ex.8.3.4: Antag att den sanna andelen bilar med felsiktade framlyktor är 0.7 (forskaren vet inte detta). Vad är då sannolikheten att vårt test inte kan upptäcka detta? Teststyrkan?

Obs! Det skulle vara bra att ha ett test med liten α och stor $1 - \beta$. Stickprovsstorleken borde vara tillräckligt stor!

Hypotestest

1) Nollhypotes H_0 (motsats till H_1) accepteras om man inte hittar tillräckligt mycket bevis för att förkasta den.

2) Mothypotes H_1 (forskningshypotes) accepteras endast om tillräckligt mycket bevis mot H_0 hittas.

3) Teststatistika är en statistika som används för att bestämma om H_0 borde förkastas.

4) Kritiska området är de numeriska värdena av teststatistikan för vilka H_0 förkastas. Kritiska området väljs så att sannolikheten att teststatistikan ligger där när H_0 är sann (sannolikheten av fel av typ I) är α. Man bestämmer α och det kritiska området på förhand.

5) Värdet av teststatistikan räknas.

6) Slutsats:

   a) Värdet av teststatistikan är på det kritiska området och man förkastar H_0 på signifikansnivå α

   b) Värdet av teststatistikan är inte på det kritiska området och man accepterar H_0 på signifikansnivå α

8.4 Signifikanstest

Signifikanstest

- Används mer och mer i stället för hypotestest (datorer).
- Man fixerar inte α och det kritiska området på förhand.
- Man väljer en teststatistika och räknar dess värde. Sedan räknar man sannolikheten att observera ett värde som är minst lika extremt som det som vi har givet att H_0 är sann. Den här sannolikheten kallas ett p-**värde**.

Ex.8.4.1: Man använder mer och mer aluminium inom bilindustri (för att minska bränslekonsumption). Med en viss bil kan man köra genomsnittligt 26 mpg (miles per gallon) och $\sigma = 5$ mpg. Man hoppas att en ny design (med mer Al) skulle öka det genomsnittliga mpg. Antag att ändringen inte påverkar σ.

Hypotestest eller signifikanstest?

- Räkna p-värdet
- Om α fixt jämför p-värdet med α och

   a) om $p \leq \alpha$, förkasta H_0

   b) om $p > \alpha$, acceptera H_0

8.5 Hypotes/signifikanstest av väntevärdet, T-test

Ex: En neurolog vill veta hur en viss narkotika påverkar reaktionstid. Hon ger narkotika till råttor och mäter deras reaktionstider efter en neurologisk stimulus. Neurologen vet att reaktionstiden utan narkotika hos en råtta är $\mu_0 = 1.2$s. Möjliga hypoteser hon skulle kunna testa är

a) Den genomsnittliga reaktionstiden skiljer sig från 1.2s. Då är $H_0 : \mu = 1.2$ och $H_1 : \mu \neq 1.2$. Det här testet är ett **dubbelsidigt test** (two-tailed test).

b) Den genomsnittliga reaktionstiden är kortare än 1.2s. Då är $H_0 : \mu \geq 1.2$ och $H_1 : \mu < 1.2$. Det här testet är ett **enkelsidigt test** (one-tailed test) och ett **left-tailed** test.

c) Den genomsnittliga reaktionstiden är längre än 1.2s. Då är $H_0 : \mu \leq 1.2$ och $H_1 : \mu > 1.2$. Det här testet är ett **enkelsidigt test** och ett **right-tailed** test.

I praktiken väljer man vanligen nollhypotesen

$$H_0 : \mu = \mu_0$$

och mothypotesen är då en av de följande

$$H_1 : \mu \neq \mu_0$$

$$H_1 : \mu < \mu_0$$

$$H_1 : \mu > \mu_0$$

(Om hypotesen $\mu = \mu_0$ kan förkastas mot hypotesen $H_1 : \mu < \mu_0$, kan också hypotesen $\mu \geq \mu_0$ förkastas, och om hypotesen $\mu = \mu_0$ kan förkastas mot hypotesen $H_1 : \mu > \mu_0$, kan också hypotesen $\mu \leq \mu_0$ förkastas.)

Ex: Man är intresserad av hur mycket duvor äter. Man undersöker 16 fåglar och mäter hur mycket mat de har i sig (i gram). Man vill testa om medelvikten (av maten) skiljer sig från 1g.

Data: 2.132 3.751 2.021 2.967 2.509 1.384 1.454 0.818 0.335 2.134 1.603 1.309 1.865 2.546 2.144 0.834

8.7 Ickeparametriska metoder

Z-testet (används för att testa hypoteser angående väntevärdet när σ är känd) och T-testet (σ inte känd) antar att stickprovet kommer från en normalfördelning eller att stickprovet är tillräckligt stort. Om man inte kan anta normalfördelning och antal observationer är litet, skulle det vara bättre att använda **ickeparametriska/parameterfria** test.

Teckentest (sign test) för medianen

När man testar hypoteser angående väntevärdet är det naturligt att använda stickprovsmedelvärdet som (en del av) teststatistikan. Om man inte vet från vilken fördelning stickprovet kommer från, vet man inte fördelningen av stickprovsmedelvärdet. Medianen är en lägesparameter liksom väntevärdet och med hjälp av den kan man hitta statistikor dess fördelning inte beror på fördelningen av stickprovet.

Man antar att stickprovet är från en kontinuerlig fördelning. Medianen M av en kontinuerlig stokastisk variabel definieras så att

$$P(X > M) = P(X < M) = \frac{1}{2},$$

dvs. att medianen är det mellersta värdet av fördelningen.

Man kan testa hypoteser angående medianen. Låt X vara en kontinuerlig stokastisk variabel med medianen M och låt $X_1, X_2, ..., X_n$ vara ett stickprov från fördelningen av X (som man inte känner). Man kan testa

$$H_0 : M = M_0 \text{ mot } H_1 : M > M_0$$

eller

$$H_0 : M = M_0 \text{ mot } H_1 : M < M_0$$

eller

$$H_0 : M = M_0 \text{ mot } H_1 : M \neq M_0$$

där M_0 är ett fixt värde.

Nu har var och en av differenserna $X_i - M_0$ sannolikheten $\frac{1}{2}$ att vara positiv, sannolikheten $\frac{1}{2}$ att vara negativ och sannolikheten 0 att vara noll. Då är, om H_0 är sann, antalet positiva differenser Q_+

$$Q_+ \sim Bin(n, \frac{1}{2})$$

och $\mathbf{E}[Q_+] = \frac{n}{2}$. Dvs. att genomsnittligt hälften av differenser är positiva och hälften negativa.

Då förkastar man H_0 om man testar mot

- $H_1 : M < M_0$ och Q_+ är för litet
- $H_1 : M > M_0$ och Q_+ är för stort
- $H_1 : M \neq M_0$ och Q_+ är för stort eller för litet.

Ex. 8.7.1: Genom att använda en standard metod är mediantiden för att göra en uppgift färdigt (vid ett löpande band) 55s. Man vill testa om en ny metod minskar mediantiden.

Teoretiskt kan differensen $X_i - M_0$ inte vara noll men nollor är möjliga i praktiken. Om det finns nollor, finns det minst två möjliga sätt att hantera dem:

1) Om antalet nollor är litet jämfört med stickprovsstorleken, kan man kasta bort dem.

2) Man kan sätta nollor som $+$-tecken eller som $-$-tecken. Man gör det så att det inte hjälper att förkasta H_0. Om man testar mot $H_1 : M < M_0$, blir nollorna $+$-tecken, om man testar mot $H_1 : M > M_0$, blir nollorna $-$-tecken och om man testar mot $H_1 : M \neq M_0$, blir nollorna $+$-tecken, om det finns färre $+$-tecken, och $-$-tecken, om det finns färre $-$-tecken.