Kapitel 10

Hypotesprövning

Vad innebär hypotesprövning?

Statistisk inferens kan utföras genom att ställa upp *hypoteser* angående en eller flera av populationens *parametrar*.

Med ett hypotestest för en parameter θ kan vi undersöka rimligheten i antagandet att värdet på parametern är $H_0: \theta=\theta_0$.

Denna *nollhypotes* ställs sedan mot en *mothypotes*. En mothypotes kan vara antingen *enkel* eller *sammansatt*.

$H_1: \theta=\theta_a$ är ett exempel på en enkel mothypotes

$H_1: \theta>\theta_0$ är ett exempel på en sammansatt mothypotes

Uppgiften blir nu att utifrån ett stickprovsresultat besluta huruvida antagandet i nollhypotesen är rimligt. Om så är fallet *accepteras* nollhypotesen och annars *förkastas* nollhypotesen.

Testfunktion och kritiskt område

En *testfunktion* är en regel som bestämmer för vilka stickprovsresultat nollhypotesen ska förkastas.

Den samling stickprov för vilka testfunktionen anger att nollhypotesen ska förkastas utgör det *kritiska området* (eller *förkastelseområdet*) som betecknas *C* (eller *RR* i boken).

Den sannolikhetsmässiga storleken på det kritiska området bestäms av *signifikansnivån α* som anger risken för ett *Typ1-fel*.

$$\alpha = \Pr\left(\text{Typ I-fel}\right) = \Pr\left[(X_1, X_2, \ldots, X_n) \in C \mid \theta = \theta_0\right]$$

Testfunktionen är en statistika. För att det enkelt ska gå att avgöra vilka stickprov som tillhör det kritiska området bör dess samplingfördelning vara enkel att hantera (åtminstone då $\theta=\theta_0$).

Exempel: Hypotesprövning angående μ vid normalfördelad population

Låt $X_1, X_2, \ldots, X_n$ vara olfsv $N(\mu,\sigma)$ med σ känd. Betrakta

H_0: $\mu=\mu_0$ som ställs mot H_1: $\mu>\mu_0$.

Det känns rimligt att förkasta nollhypotesen för stickprov där stickprovsmedelvärdet $\bar{x}$ är klart större än μ_0.

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

<u>Z är en tillsynes perfekt testfunktion</u>.

1. Z mäter avståndet mellan $\overline{X}$ och μ_0. Det verkar således rimligt att förkasta nollhypotesen då Z blir stor.

2. Z är N(0,1) då nollhypotesen är sann, dvs då $\mu=\mu_0$.

Med en signifikansnivå på $\alpha=0.05$ följer det kritiska området via

$$C = \left\{(x_1, x_2, \ldots, x_n); \overline{x} > \mu_0 + 1.645 \cdot \sigma/\sqrt{n}\right\}$$

Testets styrka

Vi söker *en metod* för att finna testfunktion och kritiskt område.

För fix signifikansnivå α finns vanligtvis ett mycket stort antal potentiella kritiska områden. Vilket är det bästa?

Definition. *Styrkefunktionen* för ett test är sannolikheten att förkasta nollhypotesen mätt för olika värden på θ, dvs

$$\text{power}\left(\theta\right) = \Pr\left[\left(X_1, X_2, \ldots, X_n\right) \in C \mid \theta\right]$$

Följaktligen gäller att $power(\theta_0) = \alpha$.

Låt θ_a vara ett värde på θ i H_1. Vi vill då förstås att $power(\theta_a)$ blir så stor som möjligt.

Bästa kritiska område och MP-test

Definition. Betrakta ett test med mothypotesen H_1*:* $\theta=\theta_a$ där C är ett kritiskt område med signifikansnivå α. C sägs vara det *bästa kritiska området* (av storlek α) om det för varje annat kritiskt område C^* (av storlek α) gäller att

$$\Pr\left[(X_1, X_2, \ldots, X_n) \in C \mid \theta = \theta_a\right] \geq \Pr\left[(X_1, X_2, \ldots, X_n) \in C^* \mid \theta = \theta_a\right]$$

Definition. Det test som använder det bästa kritiska området sägs vara ett *MP-test* (*Most Powerful*).

Hur ska vi då gå tillväga för att finna det bästa kritiska området? Låt oss successivt välja punkter/stickprov $(x_1,x_2,\ldots,x_n)$ till C. Hur bör dessa punkter/stickprov väljas?

Bästa kritiska område och MP-test

Betrakta likelihoodfunktionen $L(\theta)$. Det verkar rimligt att vi för vårt kritiska område först väljer punkter $(x_1, x_2, \ldots, x_n)$ där kvoten $L(\theta_0)/L(\theta_a)$ antar ett så litet värde som möjligt.

Fortsätt sedan att välja punkter på detta sätt tills signifikans-nivån α är ”förbrukad”.

Det bästa kritiska området bör således bestå av samtliga stickprov/punkter $(x_1, x_2, \ldots, x_n)$ för vilka det gäller att

$$\frac{L(\theta_0)}{L(\theta_a)} < k$$

där konstanten k följer av signifikansnivån α.

Neyman-Pearsons lemma

Sats (Neyman-Pearsons lemma). Betrakta ett test

$$H_0: \theta=\theta_0$$

$$H_1: \theta=\theta_a$$

Låt C vara ett kritiskt område med signifikansnivå α. Låt vidare C vara konstruerat av alla stickprovspunkter $(x_1,x_2,\ldots,x_n)$ sådana att det för någon konstant k gäller att

$$\frac{L(\theta_0)}{L(\theta_a)} < k$$

Då gäller att C är det bästa kritiska området och det test som använder C som kritiskt område är ett MP-test.

Neyman-Pearsons lemma

Bevis. Låt C och C* vara kritiska områden av storlek α där C är konstruerad enligt principen att $L(\theta_0)/L(\theta_a)<1/k$.

Definition av kritiskt område ger att

$$\alpha = \Pr\left(C \mid \theta_0\right) = \Pr\left(C^* \mid \theta_0\right)$$

Vi måste nu visa att

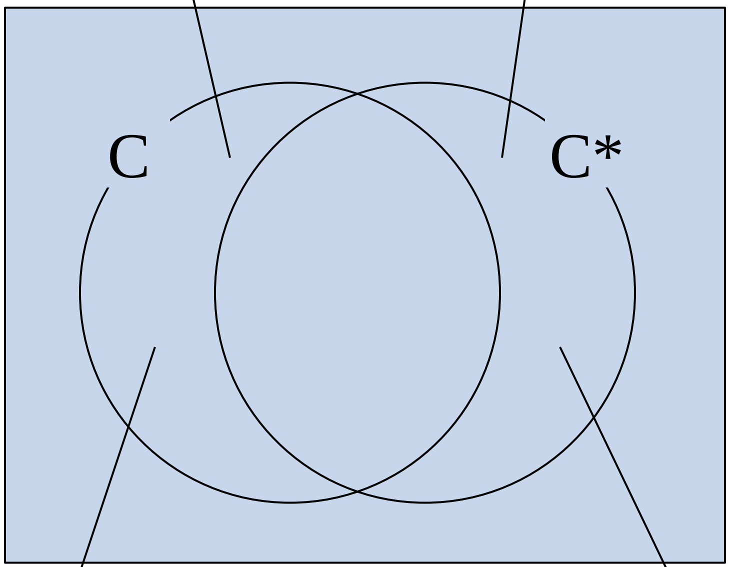
$$\Pr\left(C \mid \theta_a\right) > \Pr\left(C^* \mid \theta_a\right)$$

För att göra detta studerar vi C och C* grafiskt med ett Venn-diagram

UPPSALA UNIVERSITET

Neyman-Pearsons lemma

$$\Pr\left(C \cap \overline{C^*} \mid \theta_0\right) = \Pr\left(C^* \cap \overline{C} \mid \theta_0\right)$$



$$\Pr\left(C \cap \overline{C^*} \mid \theta_a\right) \quad > \quad k \Pr\left(C \cap \overline{C^*} \mid \theta_0\right)$$

$$\Pr\left(C^* \cap \overline{C} \mid \theta_a\right) \quad \leq \quad k \Pr\left(C^* \cap \overline{C} \mid \theta_0\right)$$

Neyman-Pearsons lemma

Detta leder till att

$$\Pr\left(C \mid \theta_a\right)$$

UPPSALA UNIVERSITET

UMP-test

Definition. Betrakta ett hypotestest angående θ där mothypotesen är en *sammansatt* hypotes.

$$H_0: \theta=\theta_0$$
$$H_1: \theta>\theta_0$$

Om det kritiska området C är det bästa kritiska området (av storlek α) för varje enkel mothypotes $H_1: \theta=\theta_a$ där $\theta_a > \theta_0$ sägs C vara ett *likformigt bästa kritiska område*. Det test som då använder C som kritiskt området sägs vara *UMP-test* (Uniformly *Most Powerful*).

Metod för att bestämma bästa kritiska område

1. Bestäm för den aktuella sannolikhetsfördelningen ett uttryck för likelihoodfunktionen $L(\theta)$.

2. Bestäm utifrån $L(\theta_0)$ och $L(\theta_a)$ ett uttryck för $L(\theta_0)/L(\theta_a) < k$.

3. Uttrycket ovan blir vanligtvis enklare att hantera efter *logaritmering*, dvs finn ett uttryck för $\ln[L(\theta_0)/L(\theta_a)] < \ln k$.

4. Det aktuella uttrycket skrivs om så att stickprovsinformationen, dvs den statistika som ska ligga till grund för testfunktionen, står i vänsterledet och en konstant (med komplicerat utseende) står i högerledet.

5. Undersök möjligheten att modifiera statistikan så att den får en hanterbar samplingfördelning. Denna modifiering blir i så fall vår testfunktion.

6. Utifrån testfunktionens samplingfördelning och signifikansnivån α bestäms ett uttryck för det bästa kritiska området. Om utseende på detta kritiska område inte beror på θ_a har vi ett UMP-test.

Uppgift 10.101

Låt $Y_1, Y_2, \ldots, Y_n$ vara oberoende och likafördelade $Exp(\theta)$.
Låt $\theta_a < \theta_0$ och betrakta hypoteserna

H_0: $\theta=\theta_0$
H_1: $\theta=\theta_a$

Bestäm bästa kritiska område med signifikansnivå α.

1. Likelihoodfunktionen för ett slumpmässigt stickprov från $Exp(\theta)$ ges av

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-y_i/\theta}$$

UPPSALA UNIVERSITET

Uppgift 10.101

2. Detta innebär att den sökta kvoten får utseendet

$$\frac{L(\theta_0)}{L(\theta_a)}$$

3. Vi logaritmerar för att få ett uttryck som är lättare att hantera.

$$\ln\left[\frac{L(\theta_0)}{L(\theta_a)}\right]$$

UPPSALA UNIVERSITET

Uppgift 10.101

4. Vi skriver nu om uttrycket så att stickprovsinformationen, dvs statistikan, står ensam i vänsterledet.

$$\sum y_i$$

$\uparrow$

Observera att olikheten får sitt utseende pga att $\theta_a < \theta_0$.

För att konstruera vårt kritiska område ska vi alltså utgå från statistikan ΣY_i.

UPPSALA UNIVERSITET

Uppgift 10.101

5. Vi betraktar statistikan ΣY_i. Då nollhypotesen är sann är dess samplingfördelning $Ga(n,\theta_0)$. Vi modifierar (se uppg 6.46)

$$W = \frac{2\sum Y_i}{\theta_0}$$

Det följer att W är χ^2-*fördelad* med $2n$ frihetsgrader då nollhypotesen är sann. Detta innebär att W utgör en utmärkt testfunktion för detta test.

6. Eftersom det kritiska område vi får via W inte beror på θ_a är det test som använder W ett UMP-test för H_a: $\theta<\theta_0$.

Tentauppgift (4:100424)

Problem. Den diskreta slumpvariabeln *X* har följande fördelning

$\mathbf{x}$	$\mathbf{0}$	$\mathbf{1}$	$\mathbf{2}$	$\mathbf{3}$
$\mathbf{p}(\mathbf{x})$	θ^2	$\theta(1-\theta)$	$\theta(1-\theta)$	$(1-\theta)^2$

Vi tänker nu testa $H_0: \theta = 0.3$ mot alternativet $H_1: \theta = 0.4$ med hjälp av endast en observation på *X*.

Vad blir det kritiska området (förkastelseområdet) för ett test med maximal styrka om den använda signifikansnivån ska vara *maximalt* 10%? Motivera.

Tentauppgift (4:100424)

Lösning. Enligt Neyman-Pearson lemma får vi maximal styrka om det kritiska området endast innehåller punkter sådana att

$$\frac{L(\theta_0)}{L(\theta_\alpha)} = \frac{L(0.3)}{L(0.4)} < k$$

där konstanten *k* bestäms av signifikansnivån. Här kan vi inte få *α=0.10…*

Det finns bara fyra tänkbara stickprovspunkter vilket innebär att vi kan ställa upp dom på följande sätt.

$\mathbf{x}$	$\mathbf{0}$	$\mathbf{1}$	$\mathbf{2}$	$\mathbf{3}$
$\mathbf{p}(\mathbf{x})$	θ^2	$\theta(1-\theta)$	$\theta(1-\theta)$	$(1-\theta)^2$
$\mathbf{L}(\mathbf{0.3})$	0.09	0.21	0.21	0.49
$\mathbf{L}(\mathbf{0.4})$	0.16	0.24	0.24	0.36
$\mathbf{L}(\mathbf{0.3})/\mathbf{L}(\mathbf{0.4})$	0.5625	0.875	0.875	1.3611

Tentauppgift (4:100424)

I och med att *maximal* signifikansnivå är *10%* följer således att

$$C = \{x : x = 0\}$$

är det enda möjliga testet och den *faktiska* signifikansnivån för testet blir därmed

$$\alpha = \Pr(C \mid \theta = 0.3) = \Pr(X = 0 \mid \theta = 0.3) = \mathbf{0.09}$$

Obs! Det finns en mer allmänt formulering av Neyman-Pearsons lemma, som låter oss konstruera en område C sådant att $\Pr(C|\theta_0) = \alpha$ med hjälp av *randomisering.*