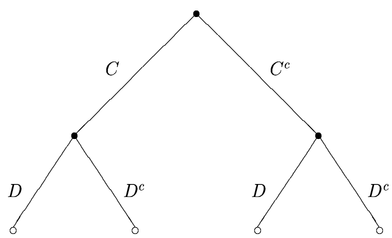
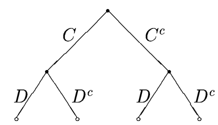
Förorenad mark. Låt y vara koncentrationen av någon giftig tungmetall. Låt x vara resultatet av en mätning av y.

Marken betraktas som förorenad om $y \geq c$, där c är ett av myndigheterna specificerat kritiskt värde. Denna händelse betecknas C.

Låt d vara den s k detekteringsnivån (obs att typiskt är $d \neq c$). Händelsen $x \geq d$ betecknas D.





Antag att du vill ”bevisa” att marken är förorenad, d.v.s att händelsen C har inträffat.

Man brukar då säga att man testar C^c mot C.

Du mäter och observerar vilken av händelserna D eller D^c som inträffar.

- Om D så har du detekterat förorening. Om du då påstår att marken är förorenad, så är risken att du har fel

  $$\alpha = P(D|C^c)$$

  Denna risk kallas testets *nivå* eller sannolikheten för ett *typ-I-fel*.

- Om D^c så har du det inte detekterat förorening. Om du då påstår att marken är ren, så är risken att du har fel

  $$\beta = P(D^c|C)$$

  Detta är sannolikheten för ett *typ-II-fel*.

- Sannolikheten för en sann detektion

  $$P(D|C) = 1 - \beta$$

  kallas för testets *upptäcktssannolikhet* eller *styrka*.

α är risken att man har bevisat en falsk hypotes

β är risken att man inte lyckats bevisa en sann hypotes

Det är svårt att samtidigt hålla båda felsannolikheterna små samtidigt

Vi ska gå igenom en standardteknik där man kan kontrollera α men inte bryr sig om β

Därför uttalar statistikern sig bara om testet ger (positivt) utslag

Om testet ej ger utslag säger han/hon inget, eftersom storleken av motsvarande felrisk β ej är känd

En statistiker uttalar sig nämligen bara då risken att han/hon tar miste är känd

Exempel 1 Antag att vi har 2000 längdskidåkare vid en tävling. Under tävlingen genomförs en dopingkontroll och 850 prövas positivt för doping. Proportionen dopade var alltså

$$\hat{p} = 0.425$$

Kan man p.g.a detta resultat påstå att färre än 50% av alla längdskidåkare dopar sig?

Exempel 2 Antag att Pontus Foxtrot vill slå vad med dig om att han med tärningen han håller i handen kommer att kasta minst 25 sexor på 100 försök. Pontus vill satsa 10 USD och undrar lite småtyket om du vågar satsa lika mycket på att han misslyckas?

Efter en snabb huvudräkning bestämmer du dig för att anta vadet.

Hur det gick? Jo, Pontus slår 28 sexor på sina 100 kast och vinner dina 10 dollar.

Fuskade Pontus eller hade han bara tur?

Rutan längst ned på sidan 326:

> The **null hypothesis**, denoted by H_0, is the claim about one or more population or process characteristics that is initially assumed to be true. The **alternative hypothesis**, denoted by H_a, is the claim that is contradictory to H_0.
>
> The null hypothesis will be rejected in favor of the alternative hypothesis only if sample evidence suggests that H_0 is false. If the sample does not strongly contradict H_0, we will continue to believe in the truth of the null hypothesis.
>
> The two possible conclusions from a hypothesis-testing analysis are then *reject* H_0 or *fail to reject* H_0.

Definition på sida 328:

> A **type I error** is the error of rejecting H_0 when H_0 is actually true.
>
> A **type II error** consists of not rejecting H_0 when H_0 is false.

Definition på sida 329:

> Maximala sannolikheten att begå ett typ-I-fel skrives α och kallas för testets *signifikansnivå*. Så att om man t.ex. väljer att testa på ”nivån” $\alpha = 0.01$, så ska risken att felaktigt förkasta en sann nollhypotes vara ≤ 0.01.
>
> Sannolikheten att göra ett typ-II-fel skrives β.
>
> $1 - \beta$ brukar man kalla testets *styrka*.

Rutan om P-värde på sidan 331:

> P-värdet eller den observerade signifikansnivån är sannolikheten för ett minst lika extremt värde som det vi erhöll (beräknad under antagandet att H_0 är sann). Ju lägre P-värdet är, desto mer osannolikare är nollhypotesen.

Exempel 8.4 på sidorna 331-332.

Recommended Daily Allowance: $\mu_0 = 15$

Data: $n = 115,\ \bar{x} = 11.3,\ s = 6.43$

Testproblem: $H_0 : \mu \geq 15$ mot $H_a : \mu < 15$

Teststatistika: $z = \dfrac{\bar{x} - 15}{s/\sqrt{n}} \overset{\text{ap}}{\approx} \mathrm{N}(0, 1)$ enl cgs

Observerat värde av z: $z = \dfrac{\bar{x} - 15}{s/\sqrt{n}} = \dfrac{11.3 - 15}{6.43/\sqrt{115}} \approx -6.17$

P-värde: $P(Z \leq -6.17) \approx 0$

Slutsats: Förkasta $H_0 : \mu \leq 15$

One-sample t-test

Förutsättning: n st oberoende N(μ, σ)-observationer $x_1, \ldots, x_n$

Beräkna $\bar{x}$, s och sedan $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

Teststatistika: $T = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$

Test av $H_0 : \mu \leq \mu_0$ mot $H_a : \mu > \mu_0$
förkasta H_0 då $P(T > t) \leq \alpha$

Test av $H_0 : \mu \geq \mu_0$ mot $H_a : \mu < \mu_0$
förkasta H_0 då $P(T < t) \leq \alpha$

Test av $H_0 : \mu = \mu_0$ mot $H_a : \mu \neq \mu_0$
förkasta H_0 då $P(T > |t|) \leq \alpha/2$

Exempel 8.5 på sida 339.

MAWL i kg

Data: 25.8, 36.6, 26.3, 21.8, 27.2
$n = 5$, $\bar{x} = 27.54$, $s = 5.47$

Testproblem: $H_0 : \mu \leq 25$ mot $H_a : \mu > 25$

Teststatistika: $T = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$

Beräkning av utfallet av T: $t = \dfrac{27.54 - 25}{5.47/\sqrt{5}} \approx 1.0$

P-värde: $P(T > t) = P(T > 1.0) = 0.187$

Slutsats: Går inte att förkasta $H_0 : \mu \leq 25$ på någon rimligt nivå α

Two-sample t-test

Förutsättning: Ett stickprov bestående av n_1 st oberoende N(μ_1, σ_1)-observationer och ett stickprov bestående av n_2 st oberoende N(μ_2, σ_2)-observationer. De två stickproven ska vara oberoende.

Data: n_1, $\bar{x}_1$, s_1 samt n_2, $\bar{x}_2$, s_2

Teststatistika: $T = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim t(\text{df})$ där

$$\text{df} \approx \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\dfrac{s_1^2/n_1}{n_1 - 1} + \dfrac{s_2^2/n_2}{n_2 - 1}}$$

Test av $H_0 : \mu_1 \leq \mu_2$ mot $H_a : \mu_1 > \mu_2$
förkasta H_0 då $P(T > t) \leq \alpha$

Test av $H_0 : \mu_1 = \mu_2$ mot $H_a : \mu_1 \neq \mu_2$
förkasta H_0 då $P(T > |t|) \leq \alpha/2$

Exempel 8.6 på sidorna 341-344.

Data 1: $n_1 = 10$, $\bar{x}_1 = 2902.8$, $s_1 = 277.3$

Data 2: $n_2 = 8$, $\bar{x}_2 = 3108.1$, $s_2 = 205.9$

Test av $H_0 : \mu_2 \leq \mu_1$ mot $H_a : \mu_2 > \mu_1$

Teststatistika: $T = \dfrac{\bar{X}_2 - \bar{X}_1}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim t(\text{df})$ där

$$\text{df} \approx \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\dfrac{s_1^2/n_1}{n_1 - 1} + \dfrac{s_2^2/n_2}{n_2 - 1}}$$

Observerade värden: $t \approx 1.8$, df $= 15.94 \approx 15$

P-värde: $P(T > 1.8) = 0.046$

Slutsats: $H_0 : \mu_2 \leq \mu_1$ kan förkastas på nivån $\alpha = 0.05$ men ej på nivån $\alpha = 0.01$.