Hyp. testing - example

1. Mom wanted to give a bag of chewing gums to her sons. Pepo really likes green ones, so she put in to his package 30 green and 10 red. Honzo likes the red ones better so his package has 30 red and 10 green. Unfortunately mom forgot to write a name on packages. She thinks that the 1st package is Pepo's, so she decided to got a handful of gums (6 gums) from the bag and test her hypothesis: If at least 3 gums will be green package has to be Pepo's.
   What is the probability of 1st and 2nd type error?

Solution :

H_0: (1st) Package belongs to Pepo.

Test: If 3 or more ch. gums (from 6) are green, then we accept the H_0, otherwise we reject it.

Number of green gums	Probability of "i" gums one handful	
	Pepo	**Honza**
i	$\mathbf{P(A}_i\mathbf{)}= \frac{\binom{30}{i}\binom{10}{6-i}}{\binom{40}{6}}$	$\mathbf{P(B}_i\mathbf{)}= \frac{\binom{30}{6-i}\binom{10}{i}}{\binom{40}{6}}$
0	0,0000547	0,154694168
1	0,001969581	0,371266003
2	0,023799103	0,321287887
3	0,126928548	0,126928548
4	0,321287887	0,023799103
5	0,371266003	0,001969581
6	0,154694168	0,0000547

Type I error : **The null hypothesis ($\mathbf{H_0}$) is true, but is rejected.**
The number of ch. gums in the handful is less then 3, but the bag belongs to Pepo.
(Warning: we look just to Pepo's column - conditional probabilities.)

$$P(i < 3|\text{Package belongs to Pepo}) = P(A_0) + P(A_1) + P(A_2) = 0,0258234$$

Type II error : **The null hypothesis is false, but fails to be rejected.**
The number of ch. gums in the handful is more then 2, but the bag belongs to Honza.
(Warning: we look just to Honza's column - conditional probabilities).

$$P(i \geq 3|\text{Package doesn't belong to Pepo}) = 1 - [P(B_0) + P(B_1) + P(B_2)] =$$
$$= P(B_3) + P(B_4) + P(B_5) + P(B_6) = 0,152751942$$

2. S kamarátom hráte kocky (Hádže sa 5x za sebou a vyhráva ten, čo má väčší súčet), vám sa moc nedarí a prvých 5 hodov boli samé 1. Potom si potrebujete odskočiť a nechce sa vám čakať, navyše kamarát tvrdí že určite vyhrá. Rozhodnete sa teda, že mu prisúdite výhru ak:

   (a) v prvom hode hodí viac ako 1,

   (b) v prvých 2 hodoch hodí aspoň raz viac ako 1,

   v opačnom prípade hypotézu zamietnete.
   Vypočítajte pravdepodobnosti chýb prvého a druhého druhu.

Solution :

H_0: Kamarát vyhrá (vy ste prehrali).

Type I error : **Nulová hypotéza (H_0) je zamietnutá, aj keď je pravdivá.**
V prvom hode padla 1 (v prvých 2 padla 1) "za podmienky", že kamarát vyhral (súčet v piatich hodoch je väčší ako 5)
(Warning: $P(A|B) = \#(A \cap B)/\#(B)$, $\omega = (x_1, ..., x_5)$.)

$$P\left(x_1 = 1 \Big| \sum x_i > 5\right) = \frac{\#\{\omega : (x_1 = 1) \cap (\sum x_i > 5)\}}{\#\{\omega : (\sum x_i > 5)\}} = \frac{6^4 - 1}{6^5 - 1} \approx 0,1666.$$

$$\left[P\left(x_1 = 1 \Big| \sum x_i > 5\right) = \frac{6^3 - 1}{6^5 - 1} \approx 0,0277\right]$$

Poznámka: $\#\{\omega : (x_1 = 1) \cap (\sum x_i > 5)\}$ = všetky možnosti kombinácie zvyšných 4 hodov mínus jediná kombinácia, kedy budú samé 1

Type II error : **Nulová hypotéza je chybná, ale my ju nezamietneme.**
V prvom hode padlo viac ako 1 (súčet prvých dvoch hodov je viac ako 2) "za podmienky", že kamarát nevyhral (súčet v piatich hodoch je 5).

$$P\left(x_1 > 1 \Big| \sum x_i = 5\right) = \frac{\#((x_1 > 1) \cap (\sum x_i = 5))}{\#(\sum x_i = 5)} = \frac{0}{1} = 0$$

$$\left[P\left(x_1 + x_2 > 2 \Big| \sum x_i = 5\right) = \frac{\#((x_1 + x_2 > 2) \cap (\sum x_i = 5))}{\#(\sum x_i = 5)} = \frac{0}{1} = 0\right]$$

3. Andrej a Bohuš majú otca, ktorý je vášnivý zberateľ. Andrej pri behu okolo regálu zhodí poličku s jednou zbierkou (7 artefaktov). Otcov systém usporiadanie tejto zbierky žiaľ nepozná a navyše každý kus je v rovnako vyzerajúcej krabici, takže zbierku na regál uloží úplne náhodne, pred bratom ale smelo prehlási, že všetko je v poriadku. Bohuš si pamätá, čo bolo v prvých 3 krabiciach a tak sa rozhodne hypotézu otestovať...
Aká je pravdepodobnosť chýb 1. a 2. druhu?

Solution :

H_0: Krabice sú zoradené správne.

Test: Správne uloženie prvých 3 krabíc .

Type I error : **Nulová hypotéza (H_0) je zamietnutá, aj keď je pravdivá.**
Hyp. zamietneme ak prvy'e 3 škatule sú nesprávne, hypotéza je pravdivá ak všetky škatule sú v správnom poradí.

$$P\left(\neg(1,2,3,x_4,...,x_7)\middle|H_0\right) = 0 \qquad (A \cap B = \emptyset)$$

Type II error : **Nulová hypotéza je chybná, ale my ju nezamietneme.**
Hyp. nezamietneme ak prvyé 3 škatule sú správne, hypotéza je nepravdivá ak aspoň jedna škatuľa je na nesprávnom mieste (Platí H_A).
(Warning: $P(A|B) = \#(A \cap B)/\#(B)$.)

$$P\left((1,2,3,x_4,...,x_7)\middle|H_A\right) = \frac{\#\{\omega : (1,2,3,x_4,...,x_7) \cap H_A\}}{\#\{\omega : (x_1,...,x_7) \neq (1,2,...,7)\}} = \frac{4!-1}{7!-1} \approx 0,00456$$

Poznámka: $\omega = (x_1,...,x_7)$ - všetky možné permutácie množiny $\{1,2,3,...,7\}$

$$\frac{\#\{\omega : (1,2,3,x_4,...,x_7) \cap H_A\}}{\#\{\omega : (x_1,...,x_7) \neq (1,2,...,7)\}} = \frac{\#\{\omega : (1,2,3,x_4,...,x_7) \cap (x_1,...,x_7) \neq (1,2,...,7))\}}{\#\{\omega : (x_1,...,x_7) \neq (1,2,...,7)\}} =$$

$$= \frac{\#\{\text{všetky možnosti na nesprávne usporiadanie 4.5.6. a 7. krabice (4 krab.)}\}}{\#\{\text{všetky možnosti na nesprávne usporiadanie všetkých 7 krabíc}\}} =$$

$$= \frac{\#\{\text{všetky možnosti usporiadania 4 krabíc mínus jediné správne}\}}{\#\{\text{všetky možnosti usporiadania 7 krabíc mínus jediné správne}\}}$$

Poznámka 2: Pokiaľ vymeníme H_0 a H_A (teda H_0 = krabice sú usporiadané nesprávne), pravdepodbnosti chýb 1. a 2. druhu budú opačne,