

Overbooking

Tomaž Martinčič

24. september 2017

1 Uvod

Znano je, da letalske družbe prodajajo več vozovnic, kot pa imajo sedežev na letalu. V angleščini za to obstajata izraza "overbooking" in "overselling". Iz izkušenj vedo, kolikšen odstotek ljudi v zadnjem trenutku odpove svoj let oziroma preprosto ne pride na vkrcanje. Da bi se izognili izgubam z letenjem s praznimi sedeži, prodajo več vozovnic, kot je prostora. Lahko pa se zgodi, da na vkrcanje pride preveč ljudi. V tem primeru so ponudniki letalskih prevozov pripravljeni plačati prostovoljcem, da prestavijo let. Včasih pa se zgodi, da šele na letalu ugotovijo, da je potnikov preveč (slika 1). In če ne najdejo prostovoljcev, letalo ne more vzleteti, kar povzroča velike denarne izgube letalski družbi. Zgodilo se je že, da so za odstranitev potnikov iz letala uporabili pretirano silo¹.

V nadaljevanju bomo prikazali statistične podatke letalskih družb, predstavili verjetnost in statistiko na kateri temelji prekomirno rezerviranje ter izpeljali formulo za izračun optimalnega števila rezervacij.

V drugem poglavju bomo predstavili verjetnost odpovedi potnika ter linearno regresijo. Prikazana sta dva grafa, ki prikazujeta trend števila potnikov ter števila zavrženih potnikov glede na četrtoletje. V 3. poglavju bomo izpeljali funkcijo za izračun optimalnega števila rezervacij glede na število sedežev na letalu. V 4. poglavju bomo predstavili interval zaupanja v povezavi s številom prostih sedežev na letalu. Zaključek pa se nahaja v 5. poglavju.



Slika 1: Na srečo prekomerno rezerviranje ne izgleda tako [2].

¹<https://www.youtube.com/watch?v=oxedJf5Cg-Y>

2 Verjetnost odpovedi potnika

Verjetnost, da bo potnik odpovedal svoj let, je odvisna od več faktorjev. Med poslovnimi uporabniki je večja verjetnost odpovedi. Lahko je sestanek na drugem koncu sveta odpovedan, ali pa se zavleče trenutni sestanek. Med turisti, ki imajo rezerviran dopust, pa je verjetnost odpovedi manjša. Ampak bodo med njimi lažje našli prostovoljce, ki bi bili pripravljeni predstaviti svoj let za nekaj ur, v zameno za denarno nadomestilo. Verjetnost odpovedi potnikov se zmanjša med prazniki (npr. za Božič), ali v času "večjih" dogodkov/tekem (Super Bowl).

Ker letalske družbe že več desetletij zbirajo podatke, lahko precej natančno napovedo, kakšna je verjetnost odpovedi potnika za recimo 14. maj, na relaciji San Francisco - New York ob sončnem vremenu. Zato se kljub naraščajočemu številu potnikov, trend zavrnjenih potnikov manjša.

United States Department of Transportation² ima na svoji spletni strani objavljene podatke za vsako četrtletje od leta 2008 naprej. Javno dostopni so podatki o številu prepeljanih potnikov (slika 2) ter številu zavrnjenih potnikov (slika 3).

Linearna regresija meri odvisnost dveh slučajnih spremenljivk (rdeča premica na sliki 2 in 3), glej [1, p. 104]. Izračunamo jo lahko po spodnji enačbi:

$$Y = \alpha + \beta X.$$

Za izračun α uporabimo:

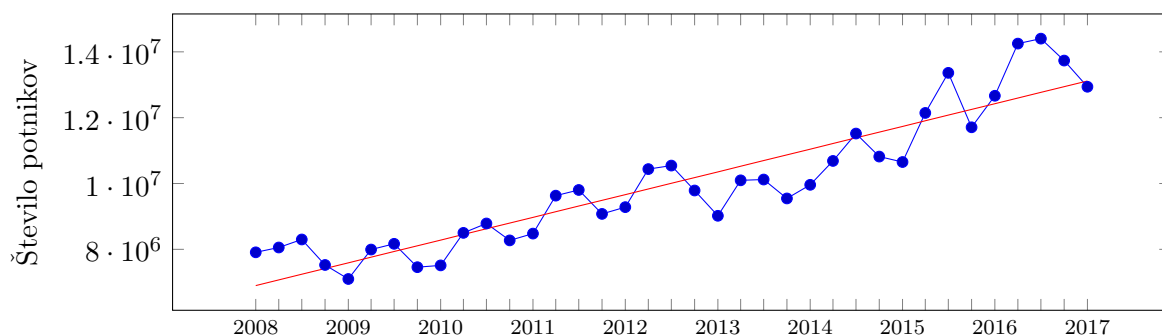
$$\alpha = y - \beta x,$$

za izračun β pa:

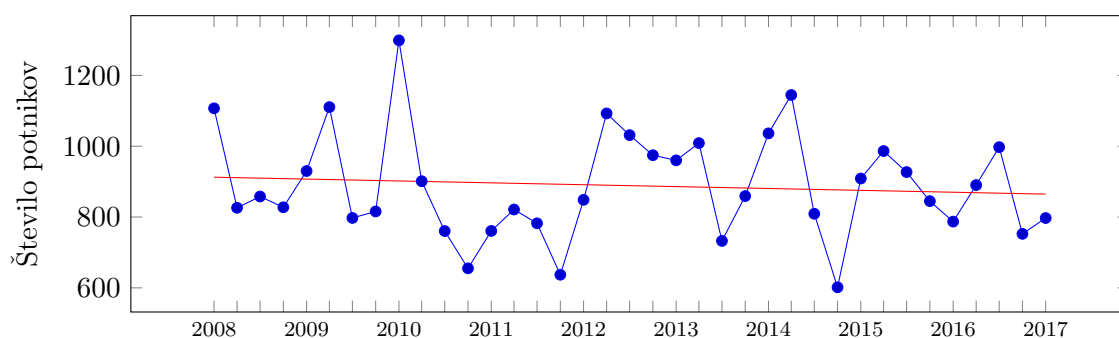
$$\begin{aligned}\beta &= \frac{k_{X,Y}}{\sigma_X^2}, \\ k_{X,Y} &= \frac{\sum_{i=1}^n (x_i - \mu_X)(y_i - \mu_Y)}{n - 1}, \\ \sigma_X^2 &= \frac{\sum_{i=1}^n (x_i - \mu_X)^2}{n - 1}, \\ \mu_X &= \frac{\sum_{i=1}^n x_i}{n}.\end{aligned}$$

S to enačbo dobimo premico, ki prikazuje odvisnost dveh slučajnih spremenljivk. V našem primeru odvisnost od četrtletja in številom potnikov.

²https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/subject_areas/airline_information/passengers_denied_confirmed_space_report/index.html



Slika 2: Število potnikov v četrtletju.



Slika 3: Število zavrženih potnikov v četrtletju.

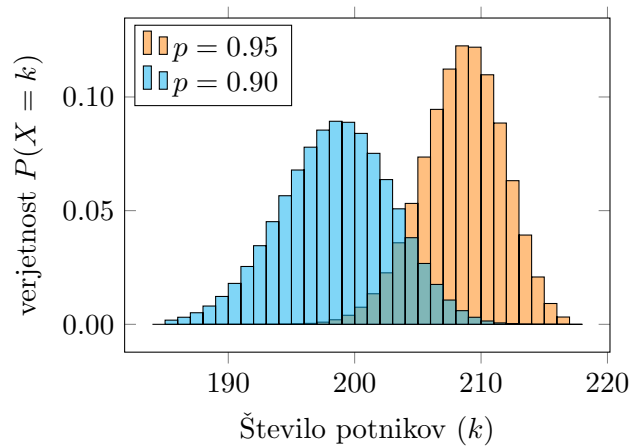
Naj bo n število rezervacij za določen let, p verjetnost, da potnik pravočasno pride na letališče in X binomsko porazdeljena slučajna spremenljivka, tj. $X \sim B(n, p)$. Verjetnost, da na letališče pravočasno pride k potnikov, izračunamo po spodnji formuli [1, s. 50]:

$$P(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}.$$

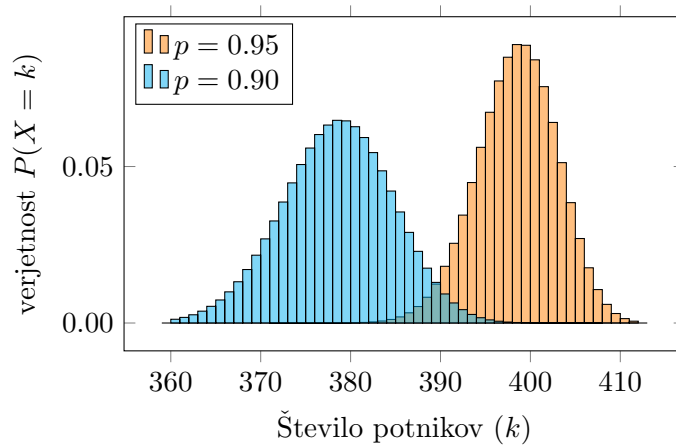
Pričakovano število potnikov izračunamo tako, da število rezerviranih kart pomnožimo z verjetnostjo, da potnik pride na rezerviran let [1, s. 51]:

$$E(X) = np.$$

Sliki 4 in 5 prikazujeta binomsko porazdelitev za verjetnost odpovedi potnika 0.1 ($p = 0.9$) in 0.05 ($p = 0.95$). Slika 4 prikazuje binomsko porazdelitev, kjer je število rezerviranih kart 220 ($n = 220$), slika 5 pa za 420 rezerviranih kart ($n = 420$).



Slika 4: Verjetnost prihoda k potnikov, če je rezerviranih 220 kart.



Slika 5: Verjetnost prihoda k potnikov, če je rezerviranih 420 kart.

V primeru s 420 rezerviranimi kartami in pri 10-% verjetnosti odpovedi potnika, je verjetnost, da pride vsaj 395 potnikov 0.218 %. To vrednost lahko izračunamo po spodnji formuli (lahko pa uporabimo aplikacijo³):

$$P(X > 395) = \sum_{i=395}^{420} p_i.$$

Lažji način za izračun približne vrednosti je uporaba centralno limitnega izreka, ki je predstavljen v naslednjem poglavju.

³<https://play.google.com/store/apps/details?id=com.mbognar.probdist>

3 Koliko kart lahko rezervirajo

Vprašanje, ki si ga zastavljajo ponudniki letalskih prevozov je, koliko kart lahko rezervirajo, da bodo z določeno verjetnostjo prepričani, da na letalu z n sedeži nihče ne bo ostal brez sedeža. Ker imajo podjetja shranjene podatke za več desetletji, lahko precej natančno napovejo, kolikšna je verjetnost p , da potnik pride na let, ki ga je rezerviral. Indikatorska slučajna spremenljivka X_i ima lahko eno izmed dveh vrednosti (0 ali 1). Vrednost 0 predstavlja, da potnik odpove let. Vrednost 1 pa, da potnik pride na let, ki ga je rezerviral:

$$X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}.$$

Povprečje, oziroma pričakovana vrednost, je enaka verjetnosti, da se potnik udeleži leta [1, s. 51]:

$$\mu = E(X_i) = p.$$

Standardni odklon izračunamo po spodnji formuli [1, s. 95]:

$$\sigma = \sqrt{D(X_i)} = \sqrt{p(1-p)}.$$

Vsota n slučajnih spremenljivk X_i se po CLI porazdeli približno normalno (za $n \geq 20$) [1, s. 106]:

$$X = X_1 + \dots + X_n = X \sim N(n\mu, \sqrt{n}\sigma).$$

Verjetnost, da bo vrednost slučajne spremenljivke X manjša od k , lahko izračunamo po spodnjih formulah [1]:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt,$$

$$P(X \leq k) = \Phi\left(\frac{n\mu - k}{\sqrt{n}\sigma}\right).$$

Vrednost $\Phi(x)$ pa lahko najdemo tudi v tabelah⁴.

Če združimo formulo za verjetnost ($P(X \leq k)$), s formulo za povprečje in formulo za standardni odklon, dobimo spodnjo formulo:

$$P(X \leq k) = \Phi\left(\frac{k - np}{\sqrt{np(1-p)}}\right).$$

Naš problem pa je malo drugačen. Izračunati želimo optimalno število rezervacij n , da bomo z verjetnostjo p' ($p' = P(X \leq k)$) prepričani, da bo dovolj sedežov (k predstavlja število sedežov). Zato moramo formulo preoblikovati:

$$\Phi^{-1}(p') > \frac{k - np}{\sqrt{np(1-p)}}$$

oziroma

$$np + \Phi^{-1}(p')\sqrt{np(1-p)} - k > 0.$$

⁴https://ucilnica.fri.uni-lj.si/pluginfile.php/3595/mod_page/content/34/Standardna_normalna_porazdelitev.pdf

Z vpeljavo nove spremenljivke $x = \sqrt{n}$, poenostavimo neenačbo v:

$$x^2 p + \Phi^{-1}(p') x \sqrt{p(1-p)} - k > 0.$$

Nato je potrebno izračunati večjo ničlo kvadratne neenačbe:

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}.$$

Dobljeno vrednost kvadriramo in zaokrožimo navzgor:

$$n = \left\lceil \left(\frac{-b + \sqrt{b^2 - 4ac}}{2a} \right)^2 \right\rceil.$$

Iz zgornjih formul lahko izpeljemo funkcijo $R(p', p, k)$, katere argumenti so opisani v tabeli 1, njena vrednost pa predstavlja število kart, ki jih lahko rezervirajo:

$$R(p', p, n) = \left\lceil \left(\frac{-\Phi^{-1}(p') \sqrt{p(1-p)} + \sqrt{\left(\Phi^{-1}(p') \sqrt{p(1-p)} \right)^2 + 4pk}}{2p} \right)^2 \right\rceil.$$

Tabela 1: Vhodni podatki in njihove zaloge vrednosti.

oznaka spremenljivke	definicijsko območje	opis
p'	$[0, 1]$	verjetnost, s katero smo prepričani, da bo na letalu dovolj sedežev
p	$[0, 1]$	verjetnost, da potnik pride na let, ki ga je rezerviral
k	$[30, \infty)$	število sedežev na letalu

4 Interval zaupanja

Z intervalom zaupanja lahko z določeno gotovostjo določimo območje, v katerem se nahaja ocenjevana vrednost. V našem primeru bomo izračunali interval, ki bo s 95-% verjetnostjo ocenil število prostih sedežev na letalu. V tabeli 2 se nahajajo podatki, spodaj pa je zapisan postopek za izračun intervala zaupanja.

Tabela 2: Število prostih sedežev na letalu (vzorci - podatki so izmišljeni).

i (številka leta)	1	2	3	4
Število prostih sedežev (x_i)	5	1	6	-1

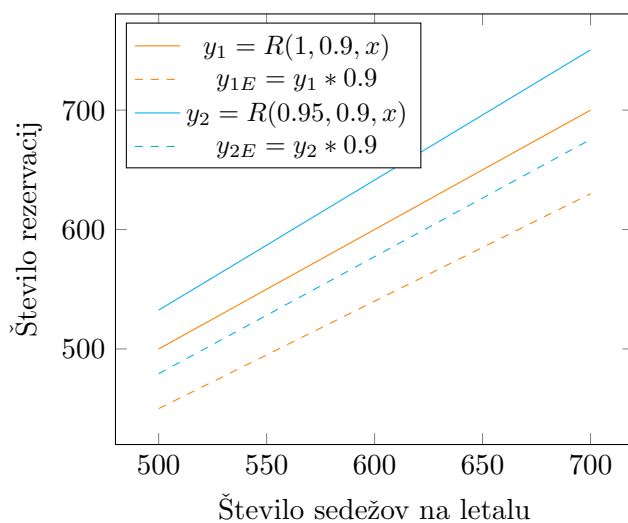
$$\begin{aligned}
 n &= 4 \\
 \bar{x} &= \frac{\sum_{i=1}^n x_i}{n} = 2.75 \\
 s &= \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = 1.5 \\
 \beta &= 0.95 \\
 \Delta &= \frac{st_{(1+\beta)/2}(n-1)}{\sqrt{n}} = 2.385 \\
 [\bar{x} - \Delta, \bar{x} + \Delta] &= [0.365, 5.135]
 \end{aligned}$$

Z verjetnostjo 0.95 lahko trdimo, da je na letalu med 0.365 in 5.135 prostih sedežev (seveda je število prostih sedežev celo število).

5 Zaključek

Na prvi pogled izgleda, da je razlog za prekomerno rezerviranje letala, zgolj želja letalske družbe po čim večjem dobičku. A v resnici imajo od tega korist tudi potniki. Glavni prednosti v očeh potnikov so cenejše karte ter možnost odpovedi rezervacije brez stroškov. Če bi prekomerno rezerviranje prepovedali, bi to podražilo letalske karte, ali pa bi jih morali plačati že ob rezervaciji.

Na sliki 6 x -os predstavlja število sedežev na letalu, y -os pa število rezerviranih kart. Polna oranžna črta prikazuje število rezerviranih kart, če želimo biti 100-% prepričani, da vsak potnik dobi sedež. Polna modra črta pa prikazuje število rezerviranih kart s 95-% verjetnostjo, da vsak potnik dobi svoj sedež. Črtkani črti pa prikazujeta pričakovano število potnikov, če je verjetnost, da podnik odpove let 10 %.



Slika 6: Število rezervacij, glede na število sedežov in verjetnost, da bodo vsi potniki dobili svoj sedež.

Če na letalu s 600 sedeži prekomerno rezerviramo karte z uporabo naše funkcije R , se pričakovan dobiček poveča za skoraj 9 %. Brez prekomirnih rezervacij, lahko pričakujemo 540 potnikov (600 rezervacij). Funkcija R pa nam priporoča, da rezerviramo 653 kart. Tako lahko pričakujemo 587.7 potnikov.

Literatura

- [1] A. Jurišić. Verjetnostni račun in statistika, skripta, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2015.
- [2] Overbooking slika.
<http://www.aviationcoaching.com/wp-content/uploads/2016/06/e54d5c7403-overbookink-OVERBOOKING.1.jpg>,
 28. 6. 2017.