

Treball final de grau
Anàlisi de l'efecte de les supersticions en el guany
del premi de la primitiva

Alumna (niu): Martínez Moran, Estela (1328691)
Tutor: Bardina i Simorra, Xavier

Grau d'Estadística Aplicada
UAB - Facultat de Ciències

9 de juny del 2016

Índex

1	Resum	4
2	Resumen	4
3	Summary	4
4	Introducció	5
4.1	Motivació	5
4.2	Dades de la BBDD	5
4.3	Validació de la base de dades	6
5	El joc de la Loteria Primitiva	7
5.1	Tipus de premi i Normativa vigent	8
6	Probabilitat del resultat de les Boles	8
6.1	Tipus de Distribució que segueixen les boles de la BBDD	11
7	El Mètode de la transformada Inversa	12
8	Anàlisi Descriptiu de la Base i Subase	14
8.1	Anàlisi Subase per a dos nombres consecutius	14
8.1.1	VR:Número d'encertants de la categoria màxima	14
8.1.2	VR:Número d'encertants per qualsevol categoria	15
8.1.3	VR:Númerod'encertants de 6, 5+C o 5	15
8.2	Anàlisi Subase per a tres nombres consecutius	17
8.2.1	VR:Número d'encertants per qualsevol categoria	17
8.2.2	VR:Número d'encertants per les categories 5,5+C i 5	17
8.2.3	VR:Número d'encertants de la categoria màxima	17
8.3	Gràfics per la base de 3 nombres consecutius	18
8.4	Anàlisi Subase per a dos nombres consecutius vs tres nombres consecutius	18
8.4.1	VR:Número d'encertants per qualsevol categoria	18
8.4.2	VR:Número d'encertants de la categoria màxima	19
8.4.3	VR:Número d'encertants per les categories 5,5+C i 6	19
8.5	Anàlisi de la subase nombres ALTS/BAIXOS	20
8.5.1	VR:Número d'encertants de la categoria màxima	21
8.5.2	VR:Número d'encertants per qualsevol categoria	21
8.5.3	VR:Número d'encertants per les categories 5,5+C i 6	22
8.5.4	Boxplot per als nombres alts/baixos segons les diferents variables resposta	22
8.6	Anàlisi de la subase nombres Parells/Senars	23
8.6.1	VR:Número d'encertants de la categoria màxima	23
8.6.2	VR:Número d'encertants per les categories 5,5+C i 6	23
8.6.3	VR:Número d'encertants per qualsevol categories	23
9	Llista dels números premiats ordenada ascendentment pel nombre mitjà d'encertants	24
10	Gràfics descriptius de les taules 4 i 5	29
10.1	Gràfic Ampliat quan hi ha bote vs no hi ha bote	29
11	Conclusions	31
11.1	Supersticions a l'hora d'apostar	31
11.2	Combinacions Favorables	31
12	Referències	34

A	Annex:Scripts	35
A.1	Teorema de la inversa	35
A.2	Anàlisi descriptiu de la base per les 3 VR	35
A.2.1	Nombres consecutius	35
A.2.2	Nombres ALTS/BAIXOS	38
A.2.3	Nombres PARELLS/SENARS	39
A.3	Llista dels números premiats ordenada ascendentment pel nombre mitja d'encertants	41

1 Resum

El present projecte del grau d'Estadística Aplicada consisteix en observar com influeixen les supersticions en el guany del joc de la Loteria Primitiva.

Segons varis estudis més del 30% [5] dels espanyols té alguna superstició o ritual a l'hora d'apostar a la loteria. Per exemple, hi han jugadors que es mantenen fidels al mateix número de loteria o que creuen que el 1234567 no té possibilitats de guanyar, quan es un nombre com qualsevol altre.

Llavors, partint de que la Primitiva és un joc d'atzar i que tots els números de les boles tenen la mateixa probabilitat de sortir, és interessant veure si la gent es deixa o no influenciar per supersticions a l'hora d'apostar. En el cas que durant l'estudi de la història de la primitiva es trobi un patró d'apostes influïdes per supersticions podrem distingir quines combinacions de nombres poden donar més premi.

Inicialment s'ha fet un recull de vàries supersticions que té la gent a l'hora d'apostar i que s'han trobat per vàries webs. Hem anat contrastant amb diverses tècniques i proves estadístiques aquestes hipòtesis.

La base que s'utilitza consta de 1.855 dades de sortejos obtinguts des del 1991 fins al 2008.

2 Resumen

El presente proyecto del grado de Estadística Aplicada consiste en observar como influyen las supersticiones en la ganancia del juego de la Lotería Primitiva.

Según varios estudios más del 30% [5] de los españoles tienen alguna superstición o ritual cuando apuestan a la lotería. Por ejemplo, hay jugadores que se mantienen fieles al mismo número de lotería o que creen que el 123 4567 no tiene posibilidades de ganar, cuando es un número cómo cualquier otro.

Entonces, partiendo de que la Primitiva es un juego de azar y que todos los números de las bolas tienen la misma probabilidad de salir, es interesante ver si la gente se deja o no influenciar por supersticiones cuando apuestan. En el supuesto de que durante el estudio de la historia de la primitiva se encuentre un patrón de apuestas influïdes por supersticiones podremos distinguir qué combinaciones de números pueden dar más premio.

Inicialmente se ha hecho una compilación de varias supersticiones que tiene la gente cuando apostamos y que se han encontrado por varias webs. Estas hipótesis se han ido contrastando con varias técnicas y pruebas estadísticas.

La base que se utiliza consta de 1.855 datos de sorteos obtenidos desde el 1991 hasta el 2008.

3 Summary

The following final project degree consists in to observation how the gain of lottery is influenced by superstitions.

According to several more studies than 30% [5] of the spanishers have any superstition or ritual when betting the lottery. For example, there are players who remain the same number of lottery or who believe that 123 4567 has no chance of wining, when is a number like any other.

Then, assuming that the Primitiva is a random game and that all numbers have the balls equally likely to leave, it's interesting to see if people are not allowed influence by superstitions when bet.

In the assumption that during the study of the early history of a pattern is found betting influences by superstitions we can distinguish which number combinations can give more prize. Firstly it made a compilation of various superstitions that people have when they are betting and which have been found for several webs. These hypotheses have been contrasted with various techniques and testing statistics. The base used consists of 1,855 dates obtained from 1991 to 2008.

4 Introducció

4.1 Motivació

La motivació principal d'aquest treball ha estat trobar una estratègia de joc a l'hora d'apostar a la primitiva ja que sempre he apostat amb resultats automàtics o influenciats per creences/supersticions. Una manera de jugar amb poca probabilitat de guany respecte a d'altres més encertades que s'explicaran més endavant.

A més de que com a estadística, tenia molta curiositat i interès en analitzar aquest món dels jocs d'atzar. En especial la primitiva, ja que és amb el que estic més familiaritzada.

En un article sobre la família "Pelayo", una família famosa per saquejar casinos que ara es dediquen a investigar sobre com obtenir combinacions guanyadores en la primitiva, es va esmentar una observació ben certa:

"Hay muchos prejuicios. Se cree que el 1234567 no tiene posibilidades de ganar, cuando es un número como cualquier otro. En este aspecto somos previsibles."

Coincidint amb la teoria dels "Pelayo", ens plantegem analitzar l'efecte de les supersticions que té la gent quan fan apostes a la primitiva.

De manera que de l'anàlisi dels números comuns que la gent aposta, podem descobrir els números favorables a obtenir un gran premi.

Ja que tots els números són equiprobables. Un principi que demostrarem més endavant.

4.2 Dades de la BBDD

Es disposa d'una base de dades amb els sortejos de la Loteria Primitiva, des del 1991 fins al 2008, extretes per via web. (Veure l'apartat de referències [6],[7] i [8]). Tot i que el joc de la primitiva va iniciar al 1985, no existia la bola de reintegrament. Es va afegir el reintegrament al juny del 1991. Per tenir una base de dades més completa hem agafat les que comencen a partir del 1991.

Variable Resposta

$$\frac{NumEncertants}{NumApostes}$$

(1)

Encertants: Número d'encertants per qualsevol categoria per cada 10000 participants.

Encertants_{sup}: Número d'encertants de la categoria màxima per cada 10 milions de participants.

Encertants_{vsup}: Número d'encertants de 6, 5+C o 5 per cada 100000 participants.

Variable explicativa

DATA,MES,ANY: Data en format "Dia/Mes/Any" de cada sorteig.

BOLA1: indicador del número de la primera bola un cop ordenades. El número mínim és 1 i el màxim 35.

BOLA2: indicador del número de la segona bola un cop ordenades. El número mínim és 2 i el màxim 39.

BOLA3: indicador del número de la tercera bola un cop ordenades. El número mínim bola és 3 i el màxim el 45.

BOLA4: indicador del número de la quarta bola un cop ordenades. El número mínim és 5 i el màxim el 47.

BOLA5: indicador del número de la cinquena bola un cop ordenades. El número mínim és 8 i el màxim el 48.

BOLA6: indicador del número de la sisena bola un cop ordenades. El número mínim és 16 i el màxim el 49.

COMPL: indicador del número complementari que ha sortit a cada sorteig. El número mínim del complementari és 1 i el màxim 49. Aquest no té ordre, un cop s'extreuen les 6 boles, s'extreu el complementari de la mateixa urna (sense repetició).

REINT: indicador del número de reintegrament. Aquest va del 0 al 9.

EUROSCAT1, . . . ,EUROSCAT5: Euros¹ recaptats en cada categoria.

ENCERCAT1, . . . ,ENCERCAT5: Número de butlletes encertants de cada categoria.

EUROSRECAPT: Recaptació total en euros de cada sorteig.

BOTEEUROS: Premi acumulat en euros².

num_{CC}: Identifica en les combinacions guanyadores els nombres consecutius.

num_{AB}: Identifica en les combinacions guanyadores els nombres alts i els baixos.

PAR: Identifica en les combinacions guanyadores els nombres parells i senars.

4.3 Validació de la base de dades

S'ha validat la base de dades per evitar inconsistències amb el suport de la pàgina web oficial on es troben tots els sortejos de l'any (veure referències a la bibliografia:[9]i[10])

Hem trobat dues dates atípiques durant la validació els dies 12/05/2001 i 20/10/2005 amb un premi acumulat molt alt. Comprovant amb la web oficial hem vist que són correctes i que com és un joc d'atzar pot ser normal trobar-se aquest tipus d'anomalies.

¹A la web obtenim pessetes i s'han convertit a euros.

² Esdevé quan cap dels sortejos de la Loteria Primitiva es queda sense encertants amb dret a premi i aquest fons que anava destinat a aquest dia, passa incrementant el concurs del dia següent, a la mateixa categoria que ha quedat buida(veure normativa).

5 El joc de la Loteria Primitiva

La Primitiva és un joc d'atzar regulat per 'Loterías y apuestas del estado(ONLAE)' que consisteix en escollir 6 números de l'1 al 49 i un número reintegrant de l'1 al 9 escollit automàticament per una màquina.

Cada aposta val 1 euro i el sortejos es realitzen els dijous i dissabtes de totes les setmanes a les 21:00 hores. Es pot jugar al sorteig diari on només participes en un sorteig, si es valida la butlleta de diumenge a dijous, es participa en el sorteig de dijous; si es valida de divendres a dissabte, es juga en el sorteig de dissabte, i al sorteig setmanal per participar en ambdós sortejos.



Figura 1: Butlleta primitiva

Existeixen 2 tipus d'apostes, la simple i la múltiple. A la simple pots jugar com a mínim 1 aposta i com a màxim 8 en una mateixa butlleta, amb 6 números per aposta. Per jugar una aposta, es marquen sis números en el primer bloc. Per jugar-ne dos, es marquen sis números en el primer bloc i sis en el segon, i així successivament. Cada aposta simple té un import d'1 euro.

Per la múltiple pots jugar més de 6 números en la teva aposta, s'ha de marcar exclusivament en el primer bloc de la butlleta. Com més gran és el nombre de múltiples seleccions, més gran serà l'import de la teva aposta.

Les apostes múltiples que es poden jugar són:

- . Casella de 44 apostes: seleccionar 5 números en el primer bloc. El conjunt format per cinc números escollits juga amb cadascun dels restants números del bloc, de manera que es formen 44 conjunts de sis pronòstics dir, 44 apostes.
- . Casella de 7 apostes: seleccionar 7 números en el primer bloc.
- . Casella de 28 apostes: seleccionar 8 números en el primer bloc.
- . Casella de 84 apostes: seleccionar 9 números en el primer bloc.
- . Casella de 210 apostes: seleccionar 10 números en el primer bloc.
- . Casella de 462 apostes: seleccionar 11 números en el primer bloc.

Cada múltiple determina el nombre d'apostes que cal pagar. A la butlleta ho trobarem a sota a l'esquerra (veure imatge butlleta).

5.1 Tipus de premi i Normativa vigent

Es destina el 55% de la recaptació a premis, que es distribueixen entre les categories següents:

- . Categoria Especial(6+R)³: si s'encerten els 6 números de la combinació guanyadora més el número del reintegrament.
- . Categoria primera: si s'encerten els 6 números de la combinació guanyadora.
- . Categoria segona(5+c): si s'encerten 5 números de la combinació guanyadora més l'extracció corresponent al número complementari⁴.
- . Categoria tercera: si s'encerten 5 números de la combinació guanyadora.
- . Categoria quarta: si s'encerten 4 números de la combinació guanyadora.
- . Categoria cinquena: si s'encerten 3 números de la combinació guanyadora.
- . Reintegrament: si el número destinat aquest premi coincideix amb el de la teva butlleta.

Durant tots aquests anys de la Loteria Primitiva hi ha hagut una sèrie de normes imposades pel 'Real Decreto 1360/1985 i publicades al BOE'. En aquestes s'explica la normativa del joc, quin fons va destinat a premis i com es reparteix, entre altres.

La distribució dels premis segons la normativa vigent, que va entrar en vigor el **dia 8 de setembre del 2012** fins a l'actualitat, és la següent:

Es destina a premis el 55 per cent de la recaudació. Un 10 per cent s'assigna a fons de premis per al reintegrament, l'altre 45 per cent es destinat a la cinquena categoria. La resta es reparteix entre les cinc primeres categories:

- . Categoria Especial: 20%
- . Categoria primera: 40%
- . Categoria segona(5+c): 6%
- . Categoria tercera: 13%
- . Categoria quarta: 21%

6 Probabilitat del resultat de les Boles

En principi les boles es recullen a partir d'un mostreig aleatori simple sense reposició, on cada bola té la mateixa probabilitat de sortir.

Sabent que en la base els números de les boles estan ordenats de menor a major, i que per tant, la bola 1 només podrà anar del 1 al 44, la segona del 2 al 45 i així successivament obtenim les següents probabilitats(veure **Taula 1. Probabilitat del resultat de les Boles**): 6

³ Nova categoria denominada Categoria Especial que entra en vigor a partir del dia 8 novembre 2012.

⁴ Nombre adicional que s'extreu en el sorteig amb l'objectiu de donar l'oportunitat de pujar el seu premi a qui ha encertat exactament cinc dels sis números de la combinació guanyadora. Aquest no es tria, és un nombre adicional que s'extreu en el sorteig i que ha de coincidir amb un dels números de la teva combinació, suposant que a més aquesta combinació inclou 5 dels 6 números de la combinació guanyadora . Per tant, només és aplicable a la 2 categoria de premis

Resultat	Probabilitat	K
Bola 1	$P(X=k) = \frac{\binom{49-k}{5}}{\binom{49}{6}}$	$k = 1, \dots, 44$
Bola 2	$P(X=k) = \frac{\binom{k-1}{1} \binom{49-k}{4}}{\binom{49}{6}}$	$k = 2, \dots, 45$
Bola 3	$P(X=k) = \frac{\binom{k-1}{2} \binom{49-k}{3}}{\binom{49}{6}}$	$k = 3, \dots, 46$
Bola 4	$P(X=k) = \frac{\binom{k-1}{3} \binom{49-k}{2}}{\binom{49}{6}}$	$k = 4, \dots, 47$
Bola 5	$P(X=k) = \frac{\binom{k-1}{4} \binom{49-k}{1}}{\binom{49}{6}}$	$k = 5, \dots, 48$
Bola 6	$P(X=k) = \frac{\binom{k-1}{5}}{\binom{49}{6}}$	$k = 6, \dots, 49$

Figura 2: Resultat de les boles

	Bola1	Bola2	Bola3	Bola4	Bola5	Bola6
1	0.12245	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.10969	0.01276	0.00000	0.00000	0.00000	0.00000
3	0.09802	0.02334	0.00109	0.00000	0.00000	0.00000
4	0.08737	0.03196	0.00304	0.00007	0.00000	0.00000
5	0.07766	0.03883	0.00568	0.00027	0.00000	0.00000
6	0.06884	0.04413	0.00883	0.00065	0.00002	0.00000
7	0.06083	0.04803	0.01231	0.00123	0.00005	0.00000
8	0.05359	0.05069	0.01601	0.00205	0.00010	0.00000
9	0.04705	0.05228	0.01978	0.00312	0.00020	0.00000
10	0.04117	0.05294	0.02353	0.00445	0.00035	0.00001
11	0.03589	0.05279	0.02715	0.00603	0.00057	0.00002
12	0.03117	0.05195	0.03056	0.00786	0.00087	0.00003
13	0.02696	0.05055	0.03370	0.00991	0.00127	0.00006
14	0.02321	0.04868	0.03651	0.01217	0.00179	0.00009
15	0.01990	0.04643	0.03894	0.01460	0.00243	0.00014
16	0.01697	0.04389	0.04097	0.01718	0.00322	0.00021
17	0.01440	0.04114	0.04256	0.01986	0.00416	0.00031
18	0.01215	0.03825	0.04372	0.02261	0.00528	0.00044
19	0.01019	0.03528	0.04442	0.02538	0.00656	0.00061
20	0.00849	0.03227	0.04468	0.02813	0.00804	0.00083
21	0.00703	0.02928	0.04451	0.03082	0.00970	0.00111
22	0.00577	0.02636	0.04393	0.03338	0.01156	0.00146
23	0.00470	0.02352	0.04295	0.03579	0.01360	0.00188
24	0.00380	0.02081	0.04161	0.03799	0.01583	0.00241
25	0.00304	0.01824	0.03995	0.03995	0.01824	0.00304
26	0.00241	0.01583	0.03799	0.04161	0.02081	0.00380
27	0.00188	0.01360	0.03579	0.04295	0.02352	0.00470
28	0.00146	0.01156	0.03338	0.04393	0.02636	0.00577
29	0.00111	0.00970	0.03082	0.04451	0.02928	0.00703
30	0.00083	0.00804	0.02813	0.04468	0.03227	0.00849
31	0.00061	0.00656	0.02538	0.04442	0.03528	0.01019
32	0.00044	0.00528	0.02261	0.04372	0.03825	0.01215
33	0.00031	0.00416	0.01986	0.04256	0.04114	0.01440
34	0.00021	0.00322	0.01718	0.04097	0.04389	0.01697
35	0.00014	0.00243	0.01460	0.03894	0.04643	0.01990
36	0.00009	0.00179	0.01217	0.03651	0.04868	0.02321
37	0.00006	0.00127	0.00991	0.03370	0.05055	0.02696
38	0.00003	0.00087	0.00786	0.03056	0.05195	0.03117
39	0.00002	0.00057	0.00603	0.02715	0.05279	0.03589
40	0.00001	0.00035	0.00445	0.02353	0.05294	0.04117
41	0.00000	0.00020	0.00312	0.01978	0.05228	0.04705
42	0.00000	0.00010	0.00205	0.01601	0.05069	0.05359
43	0.00000	0.00005	0.00123	0.01231	0.04803	0.06083
44	0.00000	0.00002	0.00065	0.00883	0.04413	0.06884
45	0.00000	0.00000	0.00027	0.00568	0.03883	0.07766
46	0.00000	0.00000	0.00007	0.00304	0.03196	0.08737
47	0.00000	0.00000	0.00000	0.00109	0.02334	0.09802
48	0.00000	0.00000	0.00000	0.00000	0.01276	0.10969
49	0.00000	0.00000	0.00000	0.00000	0.00000	0.12245

Taula 1: Probabilitat del resultat de les Boles

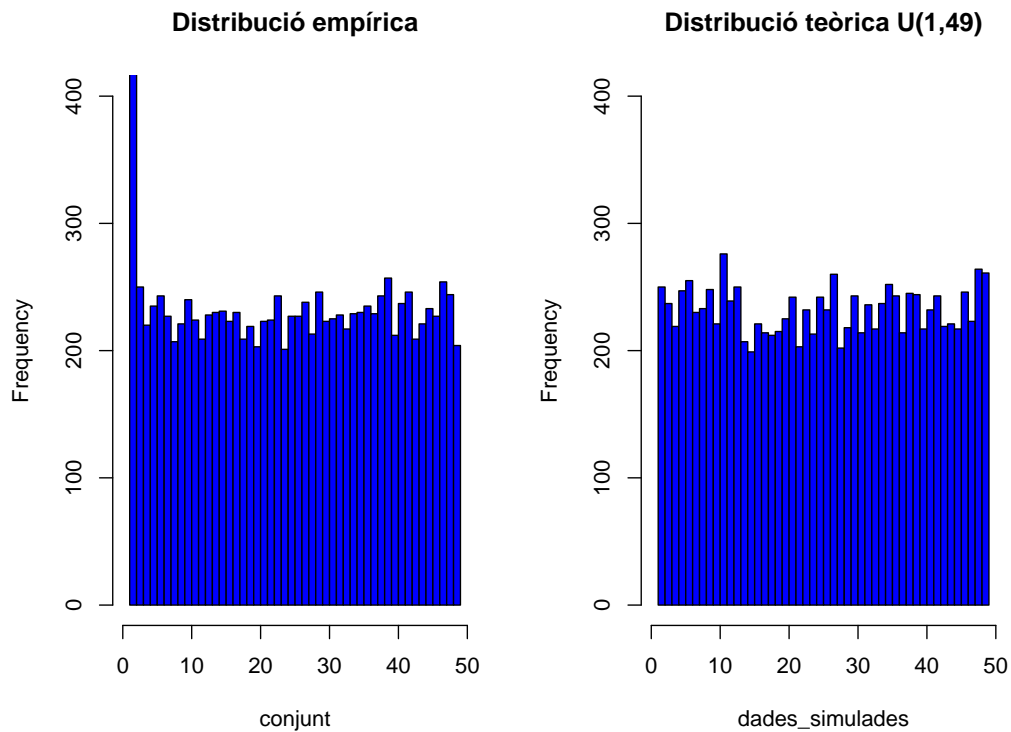
6.1 Tipus de Distribució que segueixen les boles de la BBDD

Per poder atribuir una distribució teòrica al comportament de les boles, s'ha fet un estudi utilitzant l'R amb les dades recopilades des del 1991 fins al 2008 on tenim les puntuacions ordenades de cada bola.

On es comença fent una comparació entre dos histogrames:

Un per totes les puntuacions que han resultat de la història de la loteria primitiva independentment de la bola i posició en que van sortir.

I un altre generat a partir de la distribució uniforme(1,49).



Gràficament observem que la distribució teòrica té la mateixa distribució que la empírica (utilitzem les dades de les boles en conjunt, d'aquesta manera no es té en compte ni la posició ni la bola a la que pertanyen). Les dos segueixen una uniforme(1,49).

Té sentit perquè aquesta distribució es caracteritza perquè tots els possibles valors que pot pendre la v.a. tenen la mateixa probabilitat de sortir.

7 El Mètode de la transformada Inversa

Per tal de comprovar que l'històric de les combinacions guanyadores segueix una funció de distribució uniforme(1,49) de manera més formal, utilitzarem el mètode de la transformada inversa.

Aquest mètode serveix per a la generació de nombres aleatoris de qualsevol distribució de probabilitat quan es coneix la inversa de la seva funció de distribució.

Teorema d'inversió.

Sigui X una variable aleatoria amb una funció de distribució de probabilitat acumulada F_X , invertible i contínua, i sigui F_X^{-1} la seva funció inversa(si F_X no és contínua, sigui F_X^{-1} la pseudoinversa).

Llavors, la variable aleatoria $U = F_X(X)$ té distribució uniforme en $(0;1)$.

Demostració:

$$P(F_X^{-1}(u) \leq x) = P(F_X(F_X^{-1}(u)) \leq F_X(x)) = P(u \leq F_X(x)) = \int_0^{F_X(x)} 1 dx = F_X(x)$$

$u = F_X(F_X^{-1}(u)) \sim F_X(x)$ on F_X^{-1} és la pseudoinversa.

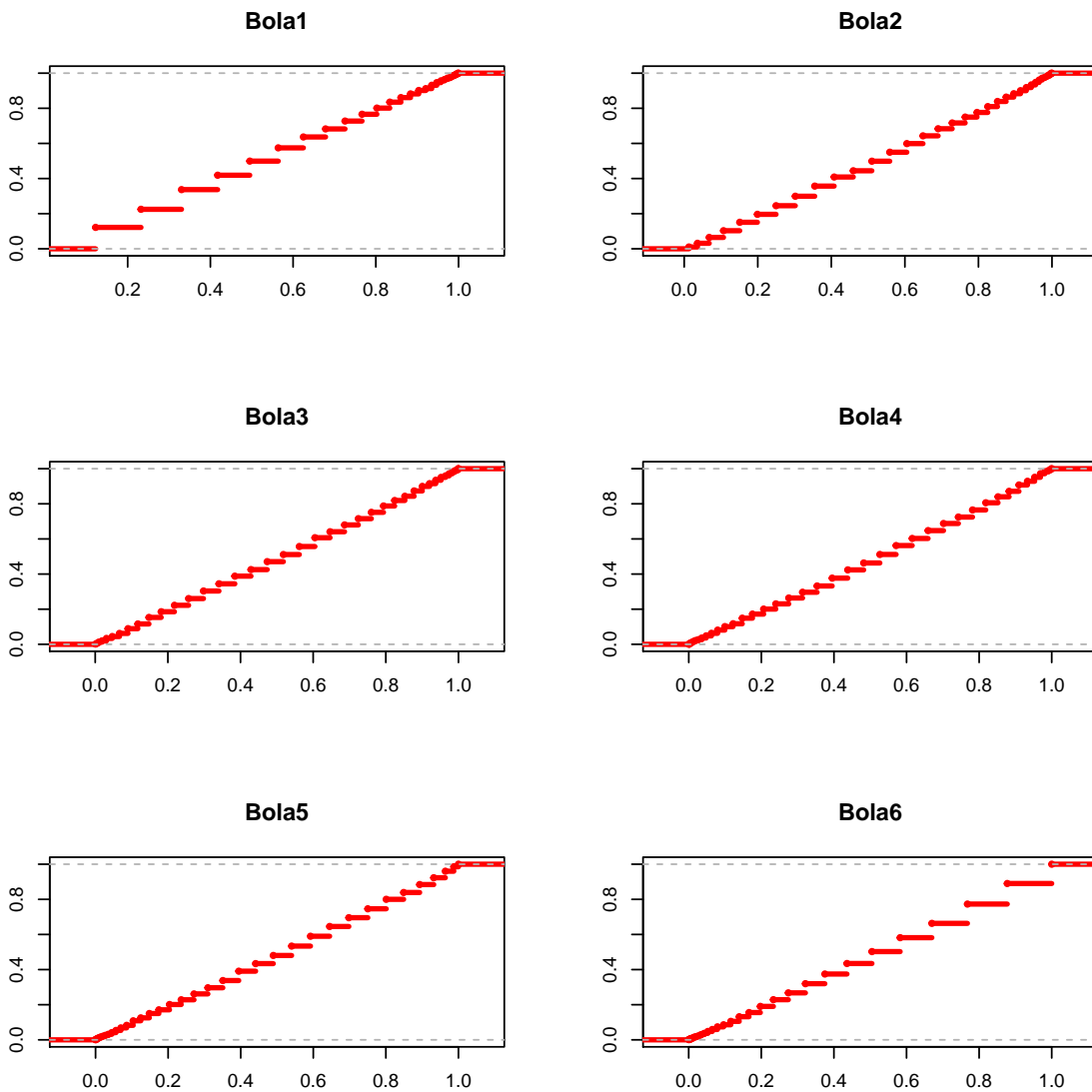
Per tant, $F_X(x) \sim Unif(0, 1)$.

Si $U \sim Unif(0, 1)$ la seva funció de distribució $F_u(x)$:

$$F_u(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in (0,1) \\ 1 & \text{si } x > 1 \end{cases}$$

Per tant, si la nostra mostra X_1, \dots, X_{1855} s'ajusta a la distribució de X , aleshores quan fem $F_X(X_1), \dots, F_X(X_{1855})$ hauria de ser una mostra d'una $Unif(0, 1)$. Fem els gràfics de la funció de distribució empírica d'aquesta nova mostra i veiem que, efectivament, s'aproxima a F_u .

Funció de distribució empírica de la mostra un cop transformada per la funció de distribució teòrica de cada bola



I efectivament observem que totes les boles tenen forma de rectes creixents i que per tant la funció empírica s'assembla a la teòrica, és a dir a una uniforme(0,1).

8 Anàlisi Descriptiu de la Base i Subbase

8.1 Anàlisi Subbase per a dos nombres consecutius

Per crear la subbase de nombres consecutius seguirem els passos següents:

1. Primer creem una nova variable y_i on no hi hauràn nombres consecutius. Per a que en un sorteig de la Loteria Primitiva no surtin dos números seguits, es necessari que existeixin almenys cinc números que s'intercalin entre les 6 boles.

Escollim 6 nombres entre 1 i 44 (no agafem les 49 boles perquè posteriorment redefinirem les variables incrementant fins a 5), aixó es pot fer de 44 sobre 6 formes diferents $\binom{44}{6}$.

Anomenem $x_1, x_2, x_3, x_4, x_5, x_6$ aquests 6 nombres escollits un cop ordenats.

Aleshores, y_i serà:

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= x_2 + 1 \\ y_3 &= x_3 + 2 \\ y_4 &= x_4 + 3 \\ y_5 &= x_5 + 4 \\ y_6 &= x_6 + 5 \end{aligned}$$

on $y_1, y_2, y_3, y_4, y_5, y_6$ són 6 nombres que estaran compresos entre 1 i 49 amb la garantia que no hi ha nombres consecutius.

2. Observem que en tota seria formada per nombres no consecutius li pots fer correspondre la respectiva sèrie al revés entre 1 i 44 restant 0, 1, 2, 3, 4, 5 al redefinir la nova variable un cop estan ordenats. I després tornant a restar la nova variable y_i amb y_{i+1} , creant un nou vector de dades.
3. Quan aquest vector de dades hi hagi un 0 li direm que són consecutius i si es $x > 0$ serà no consecutiu.

Pert tant la probabilitat que en una combinació no hi hagin 2 nombres consecutius és igual a $\binom{44}{6} \div \binom{49}{6} = 0.5048$

A continuació farem l'anàlisi dels nombres consecutius/no consecutius per les tres variables respostes:

8.1.1 VR:Número d'encertants de la categoria màxima

Dins d'aquesta categoria tenim que la mitjana del nombre d'encertants per als nombres no consecutius(1114) és superior a la de consecutius(757). Segons la taula anova següent, amb un nivell de confiança del 95%, diem que aquesta diferència es significativa:

```
##           Df      Sum Sq  Mean Sq F value   Pr(>F)
## base_cc    1 5.876e+07 58760867   45.07 2.52e-11 ***
## Residuals 1853 2.416e+09 1303767
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.1.2 VR:Número d'encertants per qualsevol categoria

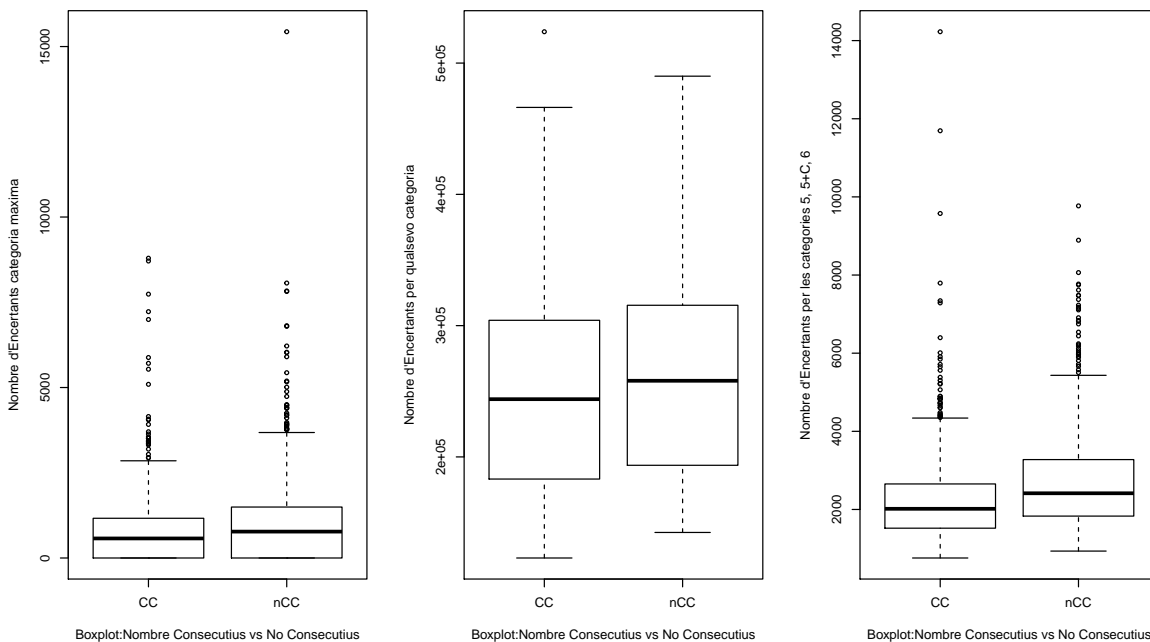
En aquesta categoria també ens trobem que el nombre mitjà d'encertants per a nombres no consecutius (259999) és superior als consecutius(249480). Segons la taula anova següent, amb un nivell de confiança del 95%, diem que aquesta diferència es significativa:

```
##           Df      Sum Sq   Mean Sq F value Pr(>F)
## base_cc      1 5.127e+10 5.127e+10   9.77 0.0018 **
## Residuals 1853 9.724e+12 5.247e+09
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.1.3 VR:Númerod'encertants de 6, 5+C o 5

En aquesta categoria ens tornem a trobar amb que el nombre mitjà d'encertants en els nombres no consecutius(2735) és major respecte als consecutius(2259). Segons la taula anova següent, amb un nivell de confiança del 95%, diem que aquesta diferència es significativa:

```
##           Df      Sum Sq   Mean Sq F value Pr(>F)
## base_cc      1 1.050e+08 105014587  72.53 <2e-16 ***
## Residuals 1853 2.683e+09  1447836
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



En el primer gràfic observem que hi ha un outlier en la mitjana del nombre d'encertants de la categoria màxima. La combinació d'aquest outlier es 2 14 22 27 34 48

En el segon gràfic podem distingir un valor atípic en la mitjana d'encertants per qualsevol categoria, que correspon a la combinació:

13 16 17 22 25 34

Per al tercer gràfic tornem a observar una dada atípica:

13 16 17 22 25 34

Amb una mitjana d'encertants igual a 14227, 11692 i 9575 respectivament.

8.2 Anàlisi Subase per a tres nombres consecutius

Per tal de trobar la proporció de 3 nombres consecutius en qualsevol combinació aleatòria sense reposició hem creat una funció.

Els passos que es segueixen en aquesta funció són:

1. Creem una nova variable(y) que sorgeix de la resta de 0, 1, 2, 3, 4, 5 a cada combinació. On aquestes combinacions són combinacions ordenades generades aleatòriament sense reposició.
2. Efectuem la diferència entre la variable y_i i y_{i+1} , creant d'aquesta manera un nou vector de dades(t).
3. Finalment per aquesta última variable efectuem la suma entre t_i i t_{i+1} .
4. Quan aquest vector de dades hi hagi un 0 li direm que són consecutius i si es $x > 0$ serà no consecutiu.

Fem una simulació de 1000000 iteracions per tal de trobar la proporció teòrica de tres nombres consecutius i ens resulta que la proporció és aproximadament igual a 0.050962.

8.2.1 VR:Número d'encertants per qualsevol categoria

En aquesta categoria tenim que la mitjana del nombre d'encertants per als nombres no consecutius(254854) és superior a la de consecutius(247005). Amb un 95% de confiança diem que aquesta diferència és no significativa:

```
##           Df      Sum Sq   Mean Sq F value Pr(>F)
## base_cc     1 4.603e+09 4.603e+09   0.874   0.35
## Residuals 1846 9.717e+12 5.264e+09
```

8.2.2 VR:Número d'encertants per les categories 5,5+C i 5

En aquesta categoria tenim que la mitjana del nombre d'encertants per als nombres no consecutius(2491) és superior a la de consecutius(2362). Amb un 95% de confiança diem que aquesta diferència és no significativa:

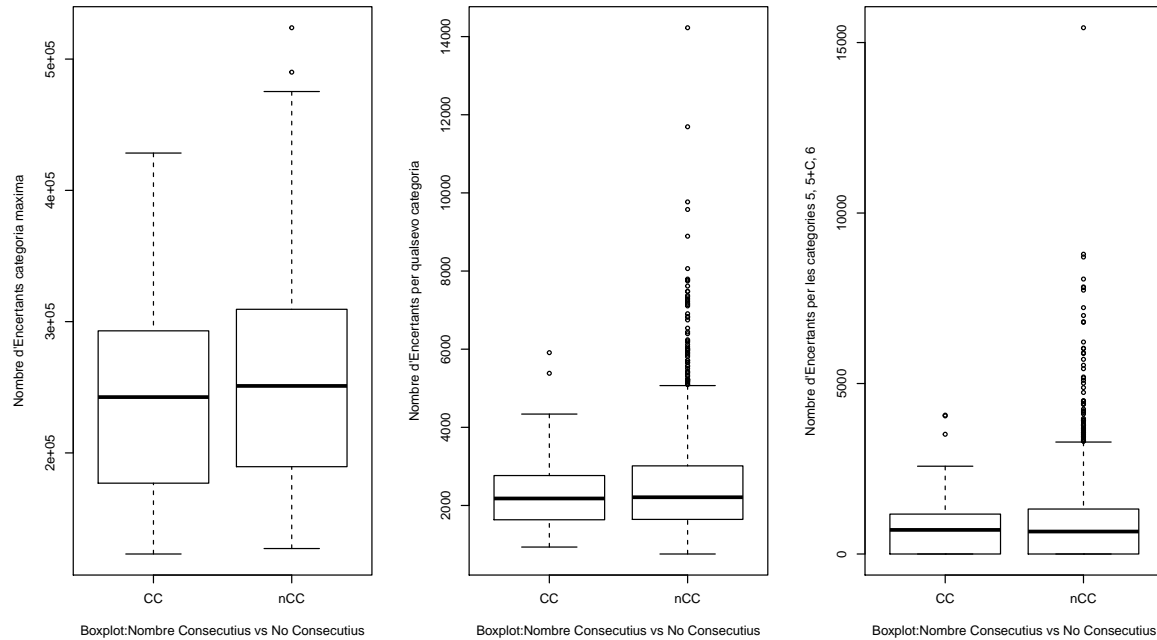
```
##           Df      Sum Sq Mean Sq F value Pr(>F)
## base_cc     1 1.247e+06 1246857   0.832   0.362
## Residuals 1846 2.767e+09 1498953
```

8.2.3 VR:Número d'encertants de la categoria màxima

En aquesta categoria també ens trobem que la mitjana d'encertants per als nombres no consecutius(936) és superior a la de consecutius(802). Segons la taula anova següent, amb un nivell de confiança del 95%, diem que aquesta diferència es no significativa:

```
##           Df      Sum Sq Mean Sq F value Pr(>F)
## base_cc    1 1.354e+06 1354148   1.012  0.315
## Residuals 1846 2.471e+09 1338440
```

8.3 Gràfics per la base de 3 nombres consecutius



Per les diferents categories les respectives combinacions atípiques/outliers són:

Encertants de qualsevol categoria: 7 14 23 26 27 28 amb una mitjana de 428414 encertants.

Encertants de les categories 5,5+C i 6: 12 14 15 16 25 26 amb una mitjana de 5911 encertants.

Encertants de la categoria superior: 1 2 3 21 23 49 amb una mitjana de 4072 encertants.

8.4 Anàlisi Subbase per a dos nombres consecutius vs tres nombres consecutius

8.4.1 VR:Número d'encertants per qualsevol categoria

La mitjana d'encertants per a 3 nombres consecutius(249532) és superior a la de per a 2 nombres consecutius(248954).

Aquesta diferència com observem en la taula anova és significativa($p\text{-value} < 0.05$).

Aquest resultat significatiu és esperable ja que en aquesta categoria nomès assegura que hi hagin com a mínim 3 encerts i infla el nombre d'encertants.

```
## [1] "data.frame"
##           Df      Sum Sq  Mean Sq F value Pr(>F)
## base_cc    1 2.588e+07 2.588e+07   0.005  0.944
## Residuals 953 5.068e+12 5.318e+09
```

8.4.2 VR:Número d'encertants de la categoria màxima

La mitjana d'encertants per a 3 nombres consecutius(753) és inferior a la de 2 nombres consecutius(805). Aquesta diferència com observem en la taula anova és no significativa(p-value>0.05).

```
## [1] "data.frame"
##           Df      Sum Sq Mean Sq F value Pr(>F)
## base_cc    1    208726   208726   0.211  0.646
## Residuals 953  944140894  990704
```

8.4.3 VR:Número d'encertants per les categories 5,5+C i 6

La mitjana d'encertants per a 3 nombres consecutius(2464) és superior a la de 2 nombres consecutius(2239). Aquesta diferència com observem en la taula anova és no significativa(p-value>0.05).

```
## [1] "data.frame"
## [1] 1
##           Df      Sum Sq Mean Sq F value Pr(>F)
## base_cc    1 3.919e+06 3918854   3.125 0.0774 .
## Residuals 953 1.195e+09 1254047
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Es sorprenent que per a 3 nombres consecutius la gent aposti més que per a 2 consecutius. Una hipòtesis és que ens trobem davant de tres tipus d'apostadors diferents:

Els supersticiosos, els aleatoris(apostes automàtiques) i els Antisupersticiosos.

Per atzar **la probabilitat d'encertar 2 nombres consecutius** es aproximadament igual al 50% front un 50% de no encertar. Seguidament **la probabilitat d'encertar 3 consecutius** es aproximadament igual al 5% front a un 95% de no encertar.

Per entendre el que volem explicar, partirem d'una mostra de 100 individus.Suposem que 55 son supersticiosos, 40 aleatoris i 5 antisupersticiosos.

Els supersticiosos, serien les persones que aposten guiant-se de supersticions, mai apostaran a nombres consecutius independentment de que siguin 2 o 3, per tant si n'hi havien 55 individus tots apostaràn als nombres no consecutius.

Els aleatoris, serien les persones que fan les apostes automàtiques, apostaran de forma aleatòria i seguirem doncs els percentatges de les probabilitats d'encertar de cada tipus de nombre consecutiu. De manera que 20 individus apostaràn nomès a nombres no consecutius, 18 individus apostaràn nomès a 2 consecutius seguits i 2 individus apostaràn nomès a 3 consecutius seguits.

Finalment, el antisupersticiosos que per dir-ho d'alguna manera serien les persones que fan anàlisis sobre la manera d'apostar dels supersticiosos i busquen apostar el contrari dels patrons trobats. De manera que d'entrada no apostaràn a cap nombre no consecutiu, 1 individu apostà nomès a 2 nombres consecutius seguits i finalment 4 individus apostaràn nomès a 3 nombres consecutius seguits (ja que serà el grup menys apostat pels supersticiosos).

El que succeeix aquí és que si sumessim cada individu per cada grup d'apostes diferent, resultaria que dels 100 individus 75% apostarien a no consecutius, 19% dels individus apostarien nomès a 2 consecutius seguits i 6

% dels individus apostarien nomès a 3 nombres consecutius seguits. On per atzar tenim que els percentatges corresponents serien 55% pels no consecutius, 40% pels 2 consecutius seguits i 5% nomès pels 3 consecutius.

De manera que observem que el grup de 3 consecutius seguits és superior en realitat que lo regit per atzar. **Aquesta idea explica el perquè hi han més apostants en el grup de 3 consecutius respecte a 2 consecutius.**

8.5 Anàlisi de la subase nombres ALTS/BAIXOS

Comprobem si realment quan al sorteig surten números alts hi ha menys encertants que quan són números baixos.

Primer busquem en quin punt parlarem de números alts o baixos en la base. Per això, sumem les 6 boles de cada un del sortejos que tenim i mirem el número mitjà.

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	48.0	128.0	150.0	150.8	173.0	244.0

Sembla simetric pero haurem de demostrar-ho teoricament ja que existeixen moltes combinacions possibles de 6 boles(13.983.816 possibilitats).

El rang de la suma total de boles estarà comprès entre la combinació mínima que serà les primeres sis boles consecutives ($1+2+3+4+5+6=21$) i la combinació màxima amb els sis últims ($44+45+46+47+48+49=279$). Quan parlem de la suma total entre les 49 boles possibles, observem que la distribució que segueix es simètrica ja que per a cada combinació possible $x_1, x_2, x_3, x_4, x_5, x_6$ hi haurà una altre combinació equivalent al mateix resultat per l'altre banda $(50 - x_1), (50 - x_2), (50 - x_3), (50 - x_4), (50 - x_5), (50 - x_6)$.

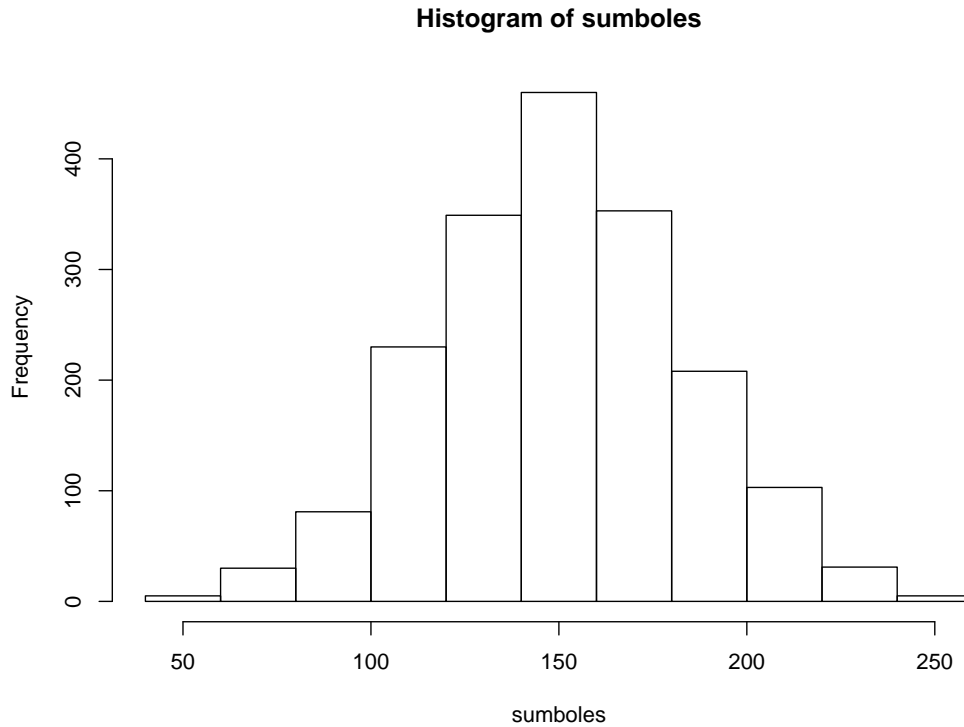
Per tant, quan fem la suma de les 13.983.816 possibilitats, és poden agrupar de dues en dues tenint en compte la seva simetria. Així doncs, quan es sumin aquestes dues combinacions equivalents sempre donarà el mateix, independentment de quines siguin $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + (50 - x_1) + (50 - x_2) + (50 - x_3) + (50 - x_4) + (50 - x_5) + (50 - x_6) = 300$.

Així doncs, podem concloure que la mitjana teòrica serà:

$$\frac{(300 \times 13983816)}{2 \cdot 13983816} = 150$$

Confirmem teòricament que el número que parteix la meitat de la suma de boles és 150.

Per tant, quan la suma sigui superior 150 es tractarà d'un sorteig d'alts i quan sigui menor de 150 es tractarà d'un sorteig de números baixos.



A continuació farem l'anàlisi dels nombres ALTS/BAIXOS per les tres variables respostes:

8.5.1 VR:Número d'encertants de la categoria màxima

La mitjana d'encertants dels nombres alts és inferior a la dels baixos(815.55 vs 1264.41 respectivament). Com observem en la següent taula anova , amb un nivell de confiança del 95%, diem que aquesta diferència és significativa:

```
##          Df    Sum Sq Mean Sq F value  Pr(>F)
## id          1 2.942e+07 29419999   20.66 6.07e-06 ***
## Residuals 1169 1.665e+09 1424285
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.5.2 VR:Número d'encertants per qualsevol categoria

La mitjana d'encertants dels nombres alts és inferior a la dels baixos(237421 vs 293368 respectivament). Com observem en la següent taula anova , amb un nivell de confiança del 95%, diem que aquesta diferència és significativa:

```
##          Df    Sum Sq  Mean Sq F value Pr(>F)
## id          1 4.571e+11 4.571e+11   99.37 <2e-16 ***
## Residuals 1169 5.377e+12 4.600e+09
## ---
```

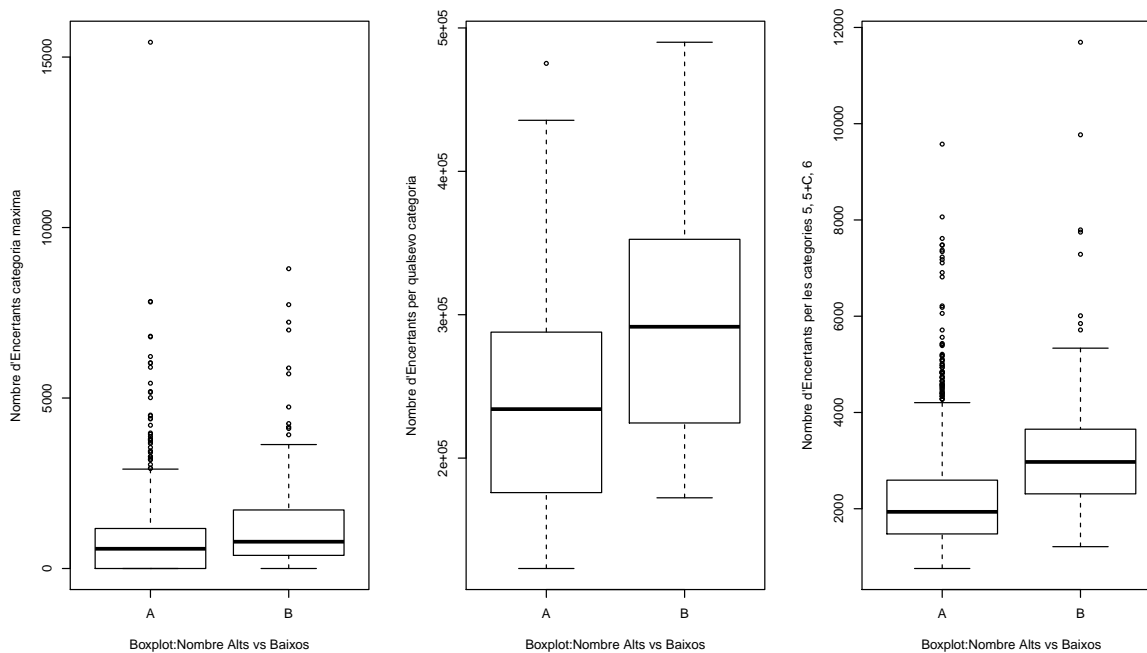
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.5.3 VR:Número d'encertants per les categories 5,5+C i 6

La mitjana d'encertants dels nombres alts és inferior a la dels baixos(2223.157 vs 3224.582 respectivament). Com observem en la següent taula anova, amb un nivell de confiança del 95%, diem que aquesta diferència és significativa:

```
## [1] 2490.072
##          Df    Sum Sq   Mean Sq F value Pr(>F)
## id       1 1.464e+08 146445516   107.6 <2e-16 ***
## Residuals 1169 1.590e+09  1360456
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.5.4 Boxplot per als nombres alts/baixos segons les diferents variables resposta



A continuació mencionarem les combinacions atípiques segons les diferents variables resposta, per ordre dels gràfics:

- Categoria màxima: 2 14 22 27 34 48
- Qualsevol categoria: 5 13 15 25 42 49
- categories 5,5+C i 6: 2 3 5 7 8 25

Amb una mitjana d'encertants igual a 15434 per a la categoria màxima, 475288 per a qualsevol categoria i 11693 per les categories 5,5+C i 6.

8.6 Anàlisi de la subbase nombres Parells/Senars

Considerem els nombre parells com aquells on a dins d'una combinació guanyadora hi han 4 o més parells. Hem eliminat els empats.

8.6.1 VR:Número d'encertants de la categoria màxima

Observem que la mitjana d'encertants de parells es superior a la de senars(968 vs 910 respectivament).Tal i com s'observa en la taula anova següent, amb un nivell de confiança del 95%, diem que aquesta diferència és NO significativa :

```
## [1] 1358 38
##           Df      Sum Sq Mean Sq F value Pr(>F)
## id          1 9.157e+05 915666 0.666 0.415
## Residuals 1356 1.864e+09 1374808
```

8.6.2 VR:Número d'encertants per les categories 5,5+C i 6

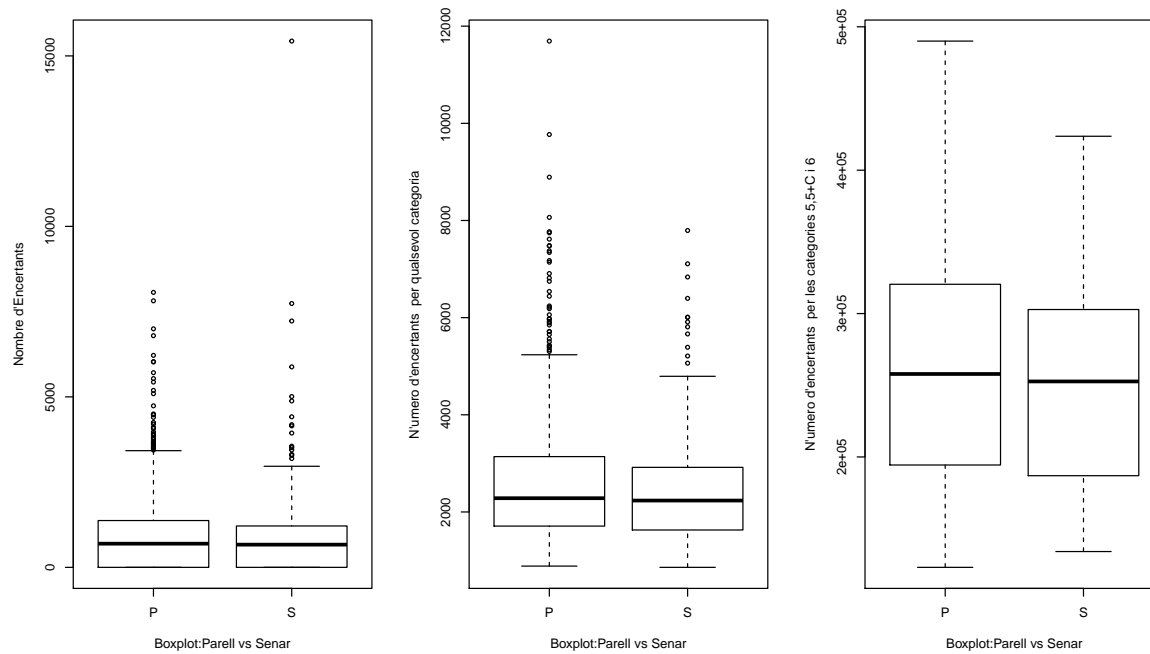
Observem que la mitjana d'encertants de parells es superior a la de senars(2604 vs 2427 respectivament).Tal i com s'observa en la taula anova següent, amb un nivell de confiança del 95%, diem que aquesta diferència és significativa :

```
## [1] 1358 38
##           Df      Sum Sq Mean Sq F value Pr(>F)
## id          1 8.170e+06 8170012 5.337 0.021 *
## Residuals 1356 2.076e+09 1530841
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.6.3 VR:Número d'encertants per qualsevol categories

Observem que la mitjana d'encertants de parells es superior a la de senars(262295 vs 249644 respectivament).Tal i com s'observa en la taula anova següent, amb un nivell de confiança del 95%, diem que aquesta diferència és significativa :

```
##           Df      Sum Sq  Mean Sq F value Pr(>F)
## id          1 4.204e+10 4.204e+10 7.697 0.00561 **
## Residuals 1356 7.406e+12 5.462e+09
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Les combinacions atípiques són les següents:

Categoria màxima: 4 7 15 17 29 33

Categories 5,5+C i 6: 1 2 4 8 9 24

Qualsevol categoria: 3 5 7 12 24 26

Amb una mitjana d'encertants igual a 8063 per la categoria màxima, 7793 per les categories 5,5+C i 6 i 490030 per qualsevol categoria.

9 Llista dels números premiats ordenada ascendentment pel nombre mitjà d'encertants

Hem creat quatre taules on en cada una es calculen les freqüències de sortir guanyador cada número(1,...,49) de cada bola (1,...,6 i complementari) i la mitjana d'encertants d'aquell número en concret, seguidament les hem ordenat ascendentment pel número mitjà d'encertants.

Dins de la categoria màxima la diferència entre les taules és que la **taula 4:Mitjana d'encertants de la categoria màxima** i **taula 5:Mitjana d'encertants de la categoria màxima quan toca bote** es selecciona de la base inicial aquells resultats on no hi ha bote i en l'altre taula no es té en compte aquest matís.

	id	freq	E(VR1)
1	25	258	288954
2	7	260	283690
3	15	260	278311
4	13	270	278149
5	5	268	273971
6	14	264	271511
7	22	267	269860
8	23	273	265292
9	12	241	264864
10	3	296	264620
11	8	253	263598
12	26	264	262580
13	6	278	261958
14	9	264	261817
15	24	249	261361
16	27	277	258970
17	33	260	258546
18	17	260	258477
19	1	266	257976
20	21	269	257571
21	11	261	256799
22	4	252	256517
23	36	278	254833
24	18	243	254379
25	16	252	253599
26	47	301	253389
27	10	279	251957
28	19	255	251795
29	45	268	250698
30	35	262	250688
31	49	243	249985
32	2	249	249909
33	28	255	249112
34	20	235	248547
35	32	281	246188
36	37	271	244912
37	43	250	244545
38	38	284	243682
39	31	259	242355
40	40	247	242158
41	42	285	240614
42	29	289	240384
43	30	259	240187
44	44	258	239795
45	34	259	239429
46	39	289	237009
47	48	287	236345
48	41	276	234537
49	46	261	232640

Taula 2: Mitjana d'encertants de totes les categories

	id	freq	E(VR2)
1	25	258	3038
2	7	260	2976
3	15	260	2908
4	13	270	2893
5	22	267	2846
6	5	268	2803
7	23	273	2785
8	14	264	2680
9	24	249	2671
10	12	241	2667
11	3	296	2654
12	27	277	2644
13	8	253	2640
14	17	260	2633
15	26	264	2608
16	4	252	2572
17	9	264	2556
18	33	260	2550
19	11	261	2535
20	36	278	2530
21	6	278	2511
22	21	269	2510
23	18	243	2492
24	2	249	2485
25	16	252	2478
26	1	266	2477
27	35	262	2445
28	19	255	2441
29	45	268	2416
30	47	301	2396
31	20	235	2378
32	10	279	2376
33	31	259	2357
34	43	250	2357
35	37	271	2353
36	32	281	2345
37	28	255	2341
38	49	243	2313
39	48	287	2286
40	42	285	2286
41	30	259	2273
42	38	284	2264
43	34	259	2262
44	40	247	2235
45	29	289	2228
46	44	258	2206
47	41	276	2162
48	39	289	2144
49	46	261	2092

Taula 3: Mitjana d'encertants de les categories 6, 5+C o 5

	id	freq	E(VR3)
1	7	192	1235
2	25	189	1209
3	15	186	1142
4	13	197	1121
5	23	201	1116
6	27	183	1102
7	4	175	1083
8	22	181	1080
9	8	183	1076
10	11	193	1070
11	24	171	1031
12	16	178	1022
13	12	157	1019
14	9	190	1017
15	26	176	1009
16	5	179	997
17	17	178	978
18	3	200	964
19	1	176	936
20	14	176	934
21	10	189	924
22	28	174	918
23	2	156	908
24	20	159	905
25	36	187	904
26	47	190	901
27	48	185	889
28	6	189	888
29	35	175	881
30	33	162	879
31	49	156	875
32	18	147	867
33	31	166	855
34	43	163	854
35	34	159	852
36	32	187	849
37	45	168	837
38	19	172	836
39	29	187	834
40	38	174	825
41	37	163	823
42	21	170	817
43	42	184	814
44	46	160	797
45	39	175	774
46	30	157	767
47	41	179	767
48	40	159	754
49	44	150	718

Taula 4: Mitjana d'encertants de la categoria maxima

	id	freq	E(VR3)
1	25	50	1636
2	8	40	1595
3	13	48	1578
4	7	48	1572
5	15	47	1500
6	12	31	1415
7	16	36	1379
8	23	40	1344
9	5	51	1316
10	22	36	1268
11	26	40	1253
12	17	36	1235
13	1	37	1214
14	47	43	1167
15	31	45	1164
16	33	37	1111
17	36	42	1094
18	3	42	1083
19	14	43	1081
20	37	41	1073
21	6	45	1072
22	24	36	1052
23	35	35	1049
24	27	39	1047
25	10	41	1046
26	49	40	1046
27	11	38	1046
28	19	51	1042
29	9	43	1038
30	38	37	1036
31	32	36	1034
32	42	44	1014
33	20	36	959
34	2	33	950
35	34	26	944
36	43	31	931
37	4	37	926
38	48	49	908
39	21	41	898
40	30	30	888
41	40	28	878
42	28	36	868
43	39	44	859
44	18	20	841
45	44	32	797
46	46	28	786
47	41	36	782
48	29	46	723
49	45	33	664

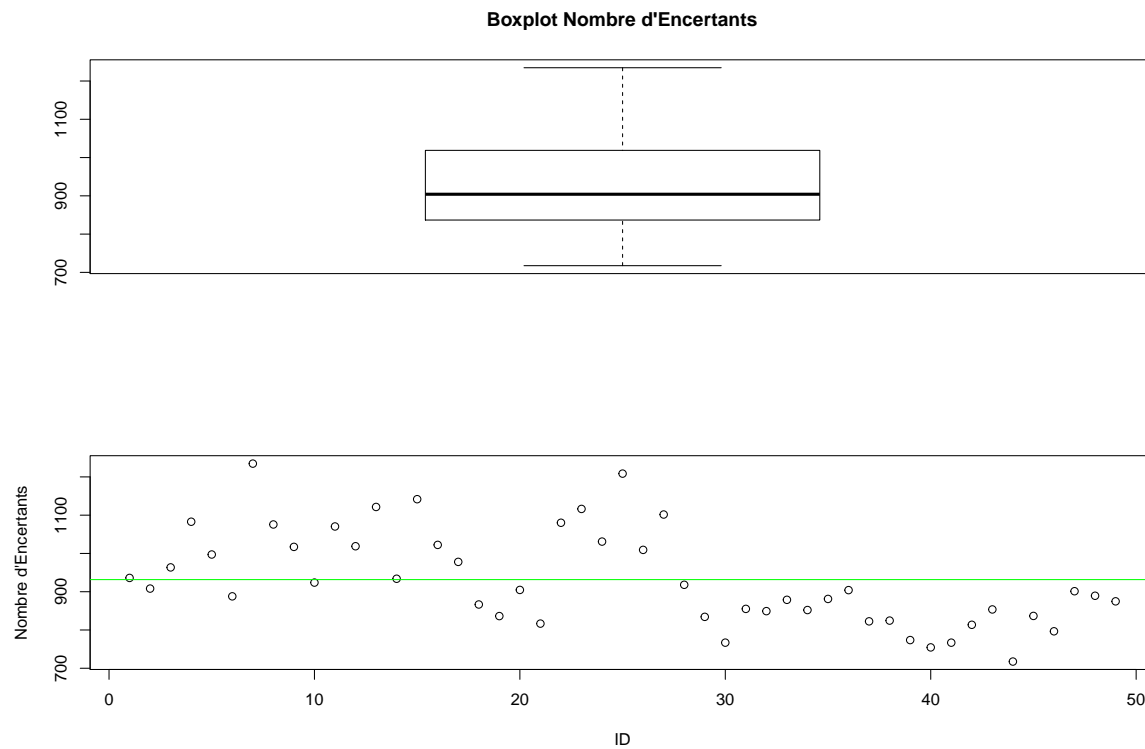
Taula 5: Mitjana d'encertants de la categoria maxima quan toca bote

10 Gràfics descriptius de les taules 4 i 5

Hem utilitzat les taules 4 i 5 referent a la mitjana d'encertants per la categoria màxima, ja que els resultats de la taula són més precisos que el de les altres variables resposta i la mida de la mostra es suficientment gran per poder treballar amb ella.

És per tant que aprofundirem en el seu anàlisi.

##	0%	25%	50%	75%	100%
##	717.5018	836.7315	904.1565	1018.8092	1234.7013

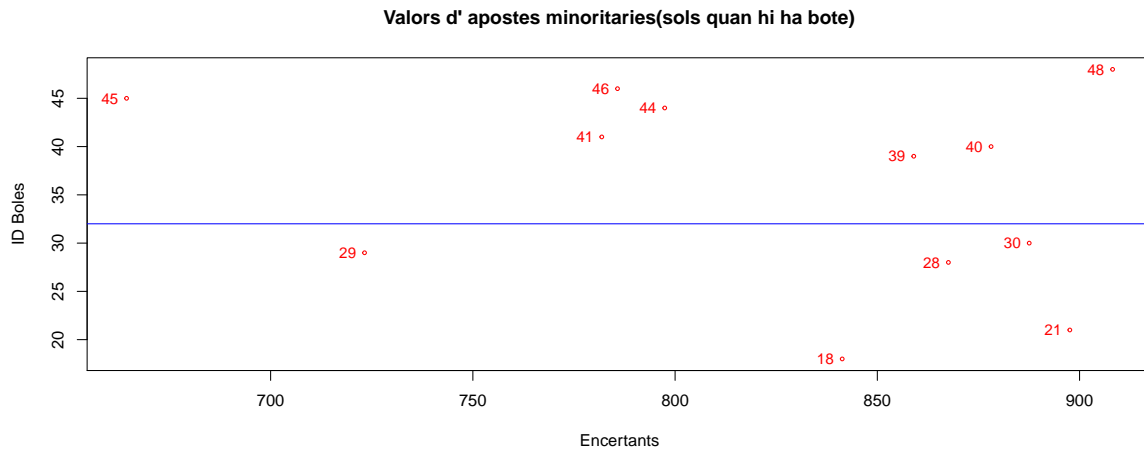
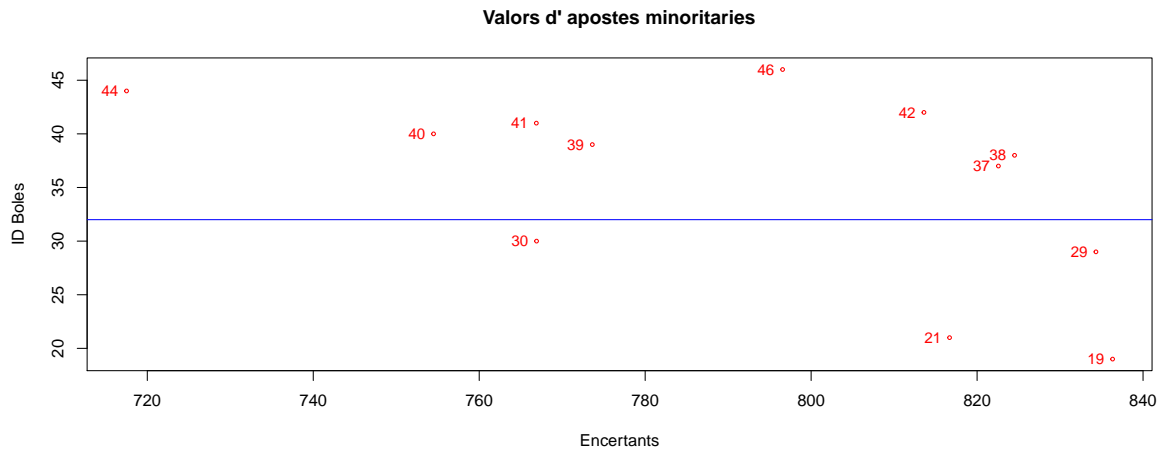


Observem una disminució del nombre d'encertants considerable a mesura de que els números de les boles són més alts.

En mitjana hi han aproximadament 905 encertants per 10 milions de jugadors.

10.1 Gràfic Ampliat quan hi ha bote vs no hi ha bote

Abans s'ha fet una descripció generals de les taules, actualment s'ha fet un ZOOM dels números als que menys aposta la gent. S'ha plotat tant per la base on sols aposten quan hi ha bote com per la base completa.



Observem en el grafic de la base general que els nombres menys apostats són en ordre ascendent:

44-40-41-30-39-46-45-21-37-38-29-19

En canvi, en el grafic de la base de les apostes quan hi ha bote, els números minoritaris per ordre ascendent són:

45-29-41-46-44-18-39-28-40-30-21-48

11 Conclusions

11.1 Supersticions a l'hora d'apostar

Molta gent tendeix a apostar a dates rellevant com ho són un aniversari o un casament. [4] Però basant-nos en el principi que hem demostrat en el punt 66.1 de que tots els números són equiprobables és favorable a l'hora d'apostar seleccionar nombres que se surten d'aquest patró d'apostes. Per tant és adient **apostar a nombres que siguin superiors al 31**.

Altres supersticions comunes són apostar a nombres de la mala sort o de la bona sort.

El número 7 [1] per exemple és un nombre molt recurrent en la nostre cultura i en algunes religions és considerat un número sagrat.

Són set els dies de la setmana, els colors de l'arc de sant martí i els pecats capitals. També es diu que un gat té set vides.

En canvi, tenim el número 13 que és considerat com el nombre de la mala sort a l'igual que el 6. **Per què?**

Una curiosa raó de per què el 13 és considerat un nombre de la mala sort és que és un número que va després del 12. El 12 és considerat un nombre perfecte. L'Any té 12 mesos, cada jornada es divideix en dia i nit en un interval de 12h de distància.

També és considerat com a nombre de la mala sort en la història occidental, es parla dels 13 assistents a l'últim sopar de Jesús, on justament Judes el va traicionar, no havia d'haver estat allà, perquè els assistents havien de ser 12.

I també es creu que van crucificar a Jesús un divendres 13 d'abril.

El número 25 en les cultures japonesa i tailandesa és considerat l'edat de la mala sort. Perque és una edat que has d'assimular moltes decisions que poden fer-te anar pel bon/mal camí. També segons la bíblia el 25 de desembre és el dia que va néixer Jesus.

I és curiós, si observem la taula4(9) i la taula5(9) que es troben en el punt 9 del treball, que els nombres influits per les supersticions que acabem de mencionar siguin els nombres més apostats, tant quan hi ha bote com quan no hi ha bote.

Però si volem que ens toqui el gran premi és millor no apostar a aquest nombres supersticiosos.

Observem en les taules que hi ha tendència també en apostar a nombres macos que siguis rodons i parells.

11.2 Combinacions Favorables

Les combinacions favorables per obtenir un gran premi segons les estadístiques obtingudes seran aquelles on la gent aposta menys.

Per tant, segons el que hem descrit al principi del treball una bona combinació es aquella que combina nombres alts, consecutius i senars.

Altres factors favorables són la distribució per desenes i apostar solament quan hi ha bote.

Per escollir nombres baixos que la gent aposta menys, ens anirem a la taula5 del punt 9 del treball9.

Considerem com a nombres baixos els que van de 1-18 (recordatori: suma nombres<150). Llavors considerem que els nombres baixos favorables són el 18, 4 i 2.

Finalment mencionar que hem utilitzat per fer les combinacions favorables la taula referent a la mitjana d'encertants per la categoria màxima.

Fixant-nos en que surti 3 nombres consecutius:

44-45-46-29-21-4
44-45-46-29-21-18
44-45-46-29-21-2

44-45-46-39-21-18
44-45-46-39-21-4
44-45-46-39-21-2

44-45-46-39-29-18
44-45-46-39-29-4
44-45-46-39-29-2

44-39-30-29-28-18
45-39-30-29-28-4
46-39-30-29-28-2

44-45-46-30-29-28

Fixant-nos en que surti 2 nombres consecutius, que segons hem vist és més favorable que no pas 3 nombres consecutius.

44-45-39-29-21-4
44-45-39-29-21-18
44-45-39-29-21-2

46-45-39-29-21-4
46-45-39-29-21-18
46-45-39-29-21-2

45-39-30-29-21-4
45-39-30-29-21-2
45-39-30-29-21-18

45-39-29-28-21-18
45-39-29-28-21-2
45-39-29-28-21-4

46-45-39-30-29-18
46-45-39-30-29-4
46-45-39-30-29-2

45-44-39-30-29-18
45-44-39-30-29-4
45-44-39-30-29-2

46-45-39-29-28-18
46-45-39-29-28-4

46-45-39-29-28-2

45-44-39-29-28-18

45-44-39-29-28-4

45-44-39-29-28-2

Aquestes combinacions junt amb les estratègies mencionades, no són una fórmula per a fer-se ric però sí per tenir més probabilitats de guanyar.

12 Referències

- [1] <https://es.wikipedia.org/wiki/Siete>
- [2] <https://es.wikipedia.org/wiki/veinticinco>
- [3] <http://www.batanga.com/curiosidades/7000/por-que-se-dice-que-el-numero-13-es-de-mala-suerte>
- [4] http://www.elconfidencial.com/tecnologia/2013-11-28/los-pelayos-contratacan-multiplicamos-por-tres-tus-opciones-de-ganar-la-primitiva_59371
- [5] <https://www.ventura24.es/blog/suerte-y-supersticion-juegan-la-loteria-de-navidad/>
- [6] <http://www.lotoideas.com/Primitiva-resultados-historicos-de-todos-los-sorteos>
- [7] <http://www.laPrimitiva.info/historico/listado.html>
- [8] <http://www.miramiprimi.miraestudio.es/historico.php>
- [9] <http://www.loteriasyapuestas.es/es/buscador?startDate=04/10/2013gameId=04type=search&endDate=04/01/2014>
- [10] <http://www.elcesardeljuego.com/2009/10/apuestas-multiples-loteria-Primitiva.html>
- [11] http://www.loteriasyapuestas.es/f/loterias/documentos/mig/estaticos/mentos_Resolution_15_2012_Modificacion_normas_La_primitiva_28912_d5785ec0.pdf
- [12] Bardina, X Càlcul de Probabilitats, Materials 139(2004)
- [13] Ross, S.M. A first course in probability. Mac Millan(1984)
- [14] Ross, S.M. A course in simulation. Macmillian Publishing Company(1990)

A Annex:Scripts

A.1 Teorema de la inversa

```
#curve(punif(x,1,49), col ="red",add = TRUE , lw =3)

datos=c(B$BOLA1,B$BOLA2,B$BOLA3,B$BOLA4,B$BOLA5,B$BOLA6)

# Ahora definimos y dibujamos la grca de la funcie distribucimpc:

#####
#Teorema de la Inversa##
#####
par(mfrow=c(3,2))
fx<-cumsum(B1)
fx2<-cumsum(B2)
fx3<-cumsum(B3)
fx4<-cumsum(B4)
fx5<-cumsum(B5)
fx6<-cumsum(B6)

ECDF=ecdf(fx[B$BOLA1])
plot(ECDF,col="red",lwd=3,main="Bola1",xlab="",ylab="")

ECDF=ecdf(fx2[B$BOLA2])
plot(ECDF,col="red",lwd=3,main="Bola2",xlab="",ylab="")

ECDF=ecdf(fx3[B$BOLA3])
plot(ECDF,col="red",lwd=3,main="Bola3",xlab="",ylab="")

ECDF=ecdf(fx4[B$BOLA4])
plot(ECDF,col="red",lwd=3,main="Bola4",xlab="",ylab="")

ECDF=ecdf(fx5[B$BOLA5])
plot(ECDF,col="red",lwd=3,main="Bola5",xlab="",ylab="")

ECDF=ecdf(fx6[B$BOLA6])
plot(ECDF,col="red",lwd=3,main="Bola6",xlab="",ylab="")
```

A.2 Anàlisi descriptiu de la base per les 3 VR

A.2.1 Nombres consecutius

```
B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

#####
##cc$Encertants_sup##
#####
```

```

cc<-subset(B,num_CC=="Consecutiu")
ncc<-subset(B,num_CC=="No Consecutiu")

#Bmax<-subset(ncc,select=c(4:10,28))
#bm<-(Bmax[order(-Bmax[,8]),])

x<-as.vector(cc$Encertants_sup)
y<-as.vector(ncc$Encertants_sup)
xy<-c(x,y)

base_cc=c(rep("CC",955),rep("nCC",900))
cc_base = data.frame(xy,base_cc)

x<-mean(cc$Encertants_sup);
y<-mean(ncc$Encertants_sup);

aov_results<-aov(xy~base_cc, data=cc_base)
summary(aov_results)

```

```

#####
##cc$Encertants#
#####

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

cc<-subset(B,num_CC=="Consecutiu")
ncc<-subset(B,num_CC=="No Consecutiu")

#bm<-subset(cc,select=c(4:10,26))
#bm<-(bm[order(-bm[,8]),])

x<-as.vector(cc$Encertants)
y<-as.vector(ncc$Encertants)
xy<-c(x,y)

base_cc=c(rep("CC",955),rep("nCC",900))
cc_base2 = data.frame(xy,base_cc)

x<-mean(cc$Encertants);
y<-mean(ncc$Encertants);

aov_results<-aov(xy~base_cc, data=cc_base2)
summary(aov_results)

```

```

#####
##cc$Encertants_usup#
#####

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

```

```

cc<-subset(B,num_CC=="Consecutiu")
ncc<-subset(B,num_CC=="No Consecutiu")

x<-as.vector(cc$Encertants_vsup)
y<-as.vector(ncc$Encertants_vsup)
xy<-c(x,y)

base_cc=c(rep("CC",955),rep("nCC",900))
cc_base3 = data.frame(xy,base_cc)

#bm<-subset(cc,select=c(4:10,27))
#bm<- (bm[order(-bm[,8]),])

x<-mean(cc$Encertants_vsup);
y<-mean(ncc$Encertants_vsup);

aov_results<-aov(xy~base_cc, data=cc_base3)
summary(aov_results)

#####
##Identificar 3 nombres consecutius##
#####

n <- 6;nsim <- 2
set.seed(19)

CC= function(n,nsim){

set.seed(19)
t<-sapply(1:nsim, function(x){return(sample(1:49,n,rep=F))})
t<-t(t)

vec=as.vector(rep(c(0:5),nsim))

m<-matrix(sort(vec),nrow=nsim,ncol=n)
tnew<-t-m
tp=apply(tnew,1,diff)

t2<-t(tp)

t3<-matrix(NA,nrow=nrow(t2),ncol=ncol(t2))
p<-0
for(i in 1:(nrow(t2))){
  for(j in 1:4){

    t3[i,j]<-t2[i,j]+t2[i,j+1]
    if(t3[i,j]==0){p[i]<-p+1;p<-p[i]}
  }
}
return(list(t,t2,t3,p,p/nsim))
}

```

```
t3<-CC(6,2)[[3]]
```

A.2.2 Nombres ALTS/BAIXOS

```
#####
#cc$Encertants_sup##
#####

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

B_A<-subset(B,num_AB=="Numeros Alts")
B_B<-subset(B,num_AB=="Numeros Baixos")
B_M<-subset(B,num_AB=="Numeros Mitjans")

#x=mean(B_A$Encertants_sup);
#y=mean(B_B$Encertants_sup);

values<-as.vector(c(B_A$Encertants_sup,B_B$Encertants_sup))
id<-c(rep("A",1000),rep("B",171))
datab<-data.frame(values,id)
#bm<-subset(B_A,select=c(4:10,27))
#bm<- (bm[order(-bm[,8]),])
#head(bm)
aov_results<-aov(values~id, data=datab)
summary(aov_results)
```

```
#####
#cc$Encertants33
#####

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

B_A<-subset(B,num_AB=="Numeros Alts")
B_B<-subset(B,num_AB=="Numeros Baixos")
B_M<-subset(B,num_AB=="Numeros Mitjans")

#x=mean(B_A$Encertants);x
#y=mean(B_B$Encertants);y

values<-as.vector(c(B_A$Encertants,B_B$Encertants))
id<-c(rep("A",1000),rep("B",171))
datab2<-data.frame(values,id)
#bm<-subset(B_A,select=c(4:10,26))
#bm<- (bm[order(-bm[,8]),])
#head(bm)
```

```

aov_results<-aov(values~id, data=datab2)
summary(aov_results)

#####
##cc$Encertants_vsup#
#####

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec
mean(B$Encertants_vsup)

B_A<-subset(B,num_AB=="Numeros Alts")
B_B<-subset(B,num_AB=="Numeros Baixos")
B_M<-subset(B,num_AB=="Numeros Mitjans")
#x=mean(B_A$Encertants_vsup);x
#y=mean(B_B$Encertants_vsup);y

values<-as.vector(c(B_A$Encertants_vsup,B_B$Encertants_vsup))
id<-c(rep("A",1000),rep("B",171))

datab3<-data.frame(values,id)
aov_results<-aov(values~id, data=datab3)
summary(aov_results)

```

A.2.3 Nombres PARELLS/SENARS

```

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

#####
##cc$Encertants_sup#
#####

B2<-subset(B,n_p!=3)
dim(B2)

B2$n_p<-ifelse(B2$n_p>=4,1,0)

Bpar<-subset(B2,B2$n_p==1)
Bsen<-subset(B2,B2$n_p==0)

x=mean(Bpar$Encertants_sup);
y=mean(Bsen$Encertants_sup);

id<-c(rep("P",1002),rep("S",356))
values<-c(as.vector(Bpar$Encertants_sup),as.vector(Bsen$Encertants_sup))
data_par<-data.frame(values,id)

anov<-aov(values~id,data=data_par)
summary(anov)

```

```

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep="," , dec

#####
##cc$Encertants_vsup#
#####

B2<-subset(B,n_p!=3)
dim(B2)

B2$n_p<-ifelse(B2$n_p>=4,1,0)

Bpar<-subset(B2,B2$n_p==1)
Bsen<-subset(B2,B2$n_p==0)

x=mean(Bpar$Encertants_vsup);
y=mean(Bsen$Encertants_vsup);

id<-c(rep("P",1002),rep("S",356))
values<-c(as.vector(Bpar$Encertants_vsup),as.vector(Bsen$Encertants_vsup))
data_par2<-data.frame(values,id)

anov<-aov(values~id,data=data_par2)
summary(anov)

```

```

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep="," , dec

#####
##cc$Encertants#
#####

B2<-subset(B,n_p!=3)

B2$n_p<-ifelse(B2$n_p>=4,1,0)

Bpar<-subset(B2,B2$n_p==1)
Bsen<-subset(B2,B2$n_p==0)

x=mean(Bpar$Encertants_vsup);
y=mean(Bsen$Encertants);

id<-c(rep("P",1002),rep("S",356))
values<-c(as.vector(Bpar$Encertants),as.vector(Bsen$Encertants))
data_par3<-data.frame(values,id)

anov<-aov(values~id,data=data_par3)
summary(anov)

```


A.3 Llista dels números premiats ordenada ascendentment pel nombre mitja d'encertants

```

B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

BZ<-subset(B,select=c(4,5,6,7,8,9,10,26))

BZ1<-melt(BZ,id=c("Encertants"))

BZ1<-BZ1[order(BZ1$value),]
##Frecuencia con que sale cada "value"(numero del 1..49) dentro de la base en formato largo

bar =as.numeric(seq(1,49, by=1))
bar.squared1.1 = rep(NA, 200000)

for(i in 1:length(bar)){
bar.squared1.1[i]<- as.data.frame(subset(BZ1,value==i))
}
bar.squared1.1=bar.squared1.1[!is.na(bar.squared1.1)]

fb1<-lapply(bar.squared1.1, function(x) length(x))
fb1<-as.numeric(fb1)

b1.1<-lapply(bar.squared1.1, function(x) mean(x))
b1.1<-as.numeric(b1.1)

x1<-cbind(bar,fb1,b1.1)
colnames(x1)<-c("id","freq","E(VR1)")
x1<-x1[order(-x1[,3]),]

x1 <- xtable(x1,caption="Mitjana d'encertants de totes les categories",digits=0)
print(x1, sanitize.text.function = function(x) {x})
B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec
BZ<-subset(B,select=c(4,5,6,7,8,9,10,27))
BZ2<-melt(BZ,id=c("Encertants_vsup"))

bar =as.numeric(seq(1,49, by=1))
bar.squared1.2 = rep(NA, 200000)
for(i in 1:length(bar)){
bar.squared1.2[i]<- as.data.frame(subset(BZ2,value==i))
}
bar.squared1.2=bar.squared1.2[!is.na(bar.squared1.2)]

fb2<-lapply(bar.squared1.2, function(x) length(x))
fb2<-as.numeric(fb2)
b1.2<-lapply(bar.squared1.2, function(x) mean(x))
b1.2<-as.numeric(b1.2)

x2<-cbind(bar,fb2,b1.2)
colnames(x2)<-c("id","freq","E(VR2)")
x2<-x2[order(-x2[,3]),]

```

```

x2 <- xtable(x2,caption="Mitjana d'encertants de les categories 6, 5+C o 5",digits=0)
print(x2, sanitize.text.function = function(x) {x})
B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec
BZ<-subset(B,select=c(4,5,6,7,8,9,10,28))
##bote:BOTEEUROS!NA->

BZ3<-melt(BZ,id=c("Encertants_sup"))
BZ3$Encertants_sup<- replace(BZ3$Encertants_sup, BZ3$Encertants_sup!=0, 1)

bar =as.numeric(seq(1,49, by=1))
bar.squared1.3 = NULL

for(i in 1:length(bar)){

bar.squared1.3[i]<- as.data.frame(subset(BZ3,value==i))

}
bar.squared1.3=as.list(bar.squared1.3[!is.na(bar.squared1.3)])

fb3<-lapply(bar.squared1.3, function(x) sum(x))
fb3<-as.numeric(fb3)
#####
B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

BZ<-subset(B,select=c(4,5,6,7,8,9,10,28))
BZ3<-melt(BZ,id=c("Encertants_sup"))

bar =as.numeric(seq(1,49, by=1))
bar.squared1.3 = NULL

for(i in 1:length(bar)){

bar.squared1.3[i]<- as.data.frame(subset(BZ3,value==i))

}
bar.squared1.3=as.list(bar.squared1.3[!is.na(bar.squared1.3)])

b1.3<-lapply(bar.squared1.3, function(x) mean(x))
b1.3<-as.numeric(b1.3)

x3<-cbind(bar,fb3,b1.3)
colnames(x3)<-c("id","freq","E(VR3)")
x3<-x3[order(-x3[,3]),]

#####
##bote:BOTEEUROS!NA###

```

```
#####

Bbote<-na.omit(B)
BZ2<-subset(Bbote,select=c(4,5,6,7,8,9,10,28))

BZ3.2<-melt(BZ2,id=c("Encertants_sup"))
BZ3.2$Encertants_sup<- replace(BZ3.2$Encertants_sup, BZ3.2$Encertants_sup!=0, 1)

bar =as.numeric(seq(1,49, by=1))
bar.squared1.4 = NULL

for(i in 1:length(bar)){

bar.squared1.4[i]<- as.data.frame(subset(BZ3.2,value==i))

}
bar.squared1.4=as.list(bar.squared1.4[!is.na(bar.squared1.4)])

fb4<-lapply(bar.squared1.4, function(x) sum(x))
fb4<-as.numeric(fb4)
#####
B<-read.csv2("C:/Users/Estela/Dropbox/TFG LOTERIA PRIMITIVA/DATf.csv", header = TRUE, sep=",", dec

BZ2<-subset(Bbote,select=c(4,5,6,7,8,9,10,28))
BZ3.2<-melt(BZ2,id=c("Encertants_sup"))

bar =as.numeric(seq(1,49, by=1))
bar.squared1.4 = NULL

for(i in 1:length(bar)){

bar.squared1.4[i]<- as.data.frame(subset(BZ3.2,value==i))

}
bar.squared1.4=as.list(bar.squared1.4[!is.na(bar.squared1.4)])

b1.4<-lapply(bar.squared1.4, function(x) mean(x))
b1.4<-as.numeric(b1.4)

x4<-cbind(bar,fb4,b1.4)
colnames(x4)<-c("id","freq","E(VR3)")
x4<-x4[order(-x4[,3]),]

id<-x3[,1]
e<-x3[,3]
```

```

par(mfrow=c(2,1))

boxplot(x3[,3],main="Boxplot Nombre d'Encertants")
m<-plot(id,e,ylab="Nombre d'Encertants",xlab="ID")
abline(m,h=mean(x3[,3]),col="green")

freq<-x3[2,]
a<-order(x3[,1])
x<-x3[a,]
freqm <- quantile(x3[,3]);freqm
x3.1<-subset(x3,x3[,3]<= 836.7315)

#####
#Grafic descriptiu#
#####

freqm1 <- quantile(x3[,3])
freqm2<- quantile(x4[,3])
x3.1<-subset(x3,x3[,3]<=836.7315)
x3.2<-subset(x4,x4[,3]<=925.8355)

par(mfrow=c(2,1))

px3.1<-plot(x3.1[,3],x3.1[,1],xlab="Encertants",ylab="ID Boles",mai="Valors d' apostes minoritaries",
text(x3.1[,3],x3.1[,1],labels=x3.1[,1], cex = 0.9, col = "red",adj=1.5)
abline(px3.1,h=32,col="blue")

px3.2<-plot(x3.2[,3],x3.2[,1],xlab="Encertants",ylab="ID Boles",mai="Valors d' apostes minoritaries",
text(x3.2[,3],x3.2[,1],labels=x3.2[,1], cex = 0.9, col = "red",adj=1.5)
abline(px3.2,h=32,col="blue")

```