

ΓΕΩΡΓΙΟΥ ΕΥΑΓΓΕΛΙΑ ΜΕΡΟΠΗ

ΣΤΑΤΙΣΤΙΚΗ ΣΥΜΠΕΡΑΣΜΑΤΟΛΟΓΙΑ

Θέμα Έρευνας

Μια εταιρεία παραγωγής ηλεκτρονικών παιχνιδιών (video games), στα πλαίσια δοκιμής ενός παιχνιδιού ρόλων φαντασίας, εξάγει εσωτερικές δοκιμές στο παιχνίδι. Ενδιαφέρεται να ερευνήσει αν το παιχνίδι είναι εξίσου δίκαιο για όλων των ειδών χαρακτήρων που θα επιλέξουν οι παίκτες, ώστε να μην υπερτερεί σημαντικά κάποιος συνδυασμός των ιδιοτήτων που δίνονται στους χαρακτήρες. Για το σκοπό αυτό συνέλεξε πληροφορία από 48 δοκιμαστές, οι οποίοι δοκίμασαν το παιχνίδι σε μια συγκεκριμένη εικονική μάχη του παιχνιδιού με τον ίδιο (τεχνητής νοημοσύνης) αντίπαλο, για τις παρακάτω μεταβλητές: (α) ζημιά στον αντίπαλο (**damage**) σε κατάλληλη μονάδα μέτρησης που έχει θέσει η εταιρεία, (β) φύλο-ηλικία του χαρακτήρα (**sex**), με κατηγορίες τις "man" (άντρας), "woman" (γυναίκα) και "children" (παιδί), (γ) φυλή του χαρακτήρα (**faction**), με κατηγορίες τις "human" (άνθρωποι), "ogre" (τέρατα) και "elf" (ξωτικά), (δ) κλάση του χαρακτήρα (**class**), με κατηγορίες τις "wizard" (μάγος) και "warrior" (μαχητής) και (ε) συμπεριφορά (**attitude**), το οποίο είναι ένα ποσοτικό χαρακτηριστικό, η τιμή του οποίου υπολογίζεται συναρτήσει κάποιων αποφάσεων που έχουν παρθεί από τον παίκτη στη διάρκεια του παιχνιδιού.

Ανάλυση

Η ομάδα που διεξήγαγε τις δοκιμές μας παρέχει ένα αρχείο data.txt με τα ευρήματά της. Αρχικά, ορίζουμε στην R το path για την φάκελο που είναι αποθηκευμένα τα data.txt και RScript μας με την εξής εντολή:

```
> setwd("C:/StatisticalInference")
```

Η ομάδα που διεξήγαγε τις δοκιμές μας παρέχει ένα αρχείο data.txt με τα ευρήματά της. Το εισάγουμε στην R με τις εξής εντολές:

```
> rm(list=ls())
```

```
> Data <- read.table(file = "data.txt", header = T, na.strings = "uns")
```

Και στη συνέχεια διαγράφουμε όλες τις παρατηρήσεις για τις οποίες δεν έχουμε τιμή η αντίστοιχο χαρακτηριστικό (δηλ. όσες έχουν NA) με την εντολή:

```
> Data <- na.omit(Data)
```

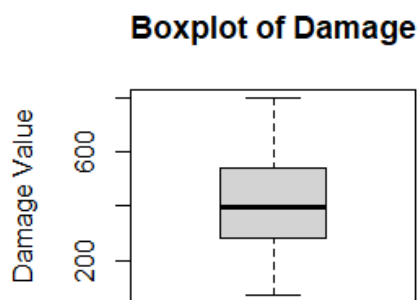
Η R επεξεργάζεται αυτή την εντολή και μας δίνει το πλαίσιο δεδομένων Data το οποίο είναι ένα data.frame και φαίνεται παρακάτω.

	damage	sex	faction	class	attitude					
1	277.48157	man	human	wizard	78.13099	24	299.39612	woman	ogre warrior	84.86424
2	556.91485	children	human	wizard	160.91966	25	196.96243	children	ogre warrior	53.61770
3	341.55235	woman	human	wizard	96.39030	26	608.12428	woman	ogre warrior	173.57687
4	543.94284	woman	ogre	warrior	155.81111	27	289.07176	woman	elf warrior	79.90764
5	418.10003	children	ogre	wizard	117.76982	28	420.22759	woman	ogre warrior	119.47224
6	496.49359	woman	ogre	wizard	143.81055	29	497.38974	children	ogre warrior	142.08133
7	727.44747	woman	human	wizard	214.06633	30	162.62395	woman	ogre warrior	40.87574
8	687.88652	children	human	wizard	201.52597	31	280.67337	woman	human warrior	76.39351
9	739.92653	woman	ogre	wizard	214.78265	32	399.27635	man	ogre wizard	114.49457
10	297.00122	woman	ogre	wizard	82.21279	33	721.87263	children	ogre wizard	210.26646
11	74.91456	man	elf	wizard	13.63325	34	377.40866	man	elf wizard	105.00977
12	383.35787	woman	ogre	wizard	110.81141	36	282.21515	woman	ogre wizard	78.38405
13	215.72197	woman	human	wizard	57.07777	37	398.26512	man	ogre warrior	111.06528
14	513.07428	man	elf	wizard	147.39881	38	265.27044	woman	elf wizard	71.36514
15	531.87218	children	elf	wizard	152.21807	39	188.36572	children	ogre warrior	47.84708
16	513.82130	man	ogre	wizard	148.43328	41	711.84000	woman	elf wizard	208.99889
17	311.94881	woman	elf	warrior	83.80284	42	105.25878	children	ogre warrior	23.01371
18	563.62944	woman	elf	warrior	159.94769	43	427.83549	children	elf warrior	119.53519
19	431.35136	children	ogre	warrior	124.46878	44	282.46687	children	human wizard	78.03080
20	540.91727	children	human	warrior	156.27282	45	378.89452	children	human warrior	107.93233
21	446.61040	man	ogre	warrior	127.76251	46	796.87913	children	elf wizard	224.69405
22	328.02945	children	ogre	wizard	92.16626	47	551.79285	man	ogre warrior	159.24057
23	232.55587	woman	elf	wizard	61.81059	48	273.40465	woman	human warrior	74.86557

Ένας από τους δημιουργούς του παιχνιδιού υποστηρίζει ότι στο συγκεκριμένο στάδιο του παιχνιδιού που διεξάγεται αυτή η μάχη, η μέση τιμή της ζημιάς που προκαλεί κάποιος παίχτης είναι μεγαλύτερη από 120 μονάδες.

Γράφοντας τις ακόλουθες εντολές παίρνουμε το Θηκόγραμμα, Ιστόγραμμα για το damage καθώς και τις 5 σημαντικές τιμές του για να έχουμε μια εικόνα για τα δεδομένα μας.

```
> par(mfrow=c(1,1))
> boxplot(Data[,1],ylab="Damage Value",main="Boxplot of Damage")
> fivenum(Data[,1])
[1] 74.91456 282.21515 398.77074 540.91727 796.87913
```

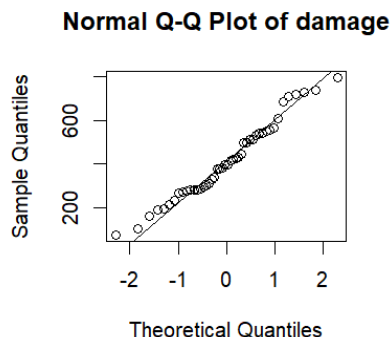
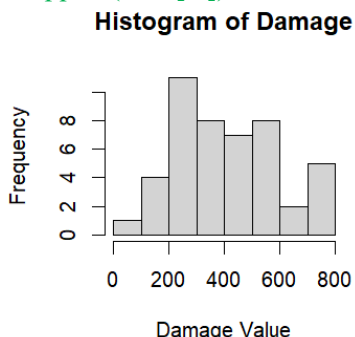


Από το θηκόγραμμα φαίνεται η υπόθεση του δημιουργού να είναι σωστή, αφού το Q1 είναι ήδη μεγαλύτερο από το 120 και η διάμεσος είναι κοντά στο 400. Μένει να κάνουμε τους κατάλληλους ελέγχους για να είμαστε βέβαιοι.

Τώρα που έχουμε μία ιδέα, μπορούμε να προχωρήσουμε στον πραγματικό μας έλεγχο. Αρχικά ελέγχουμε αν για το damage έχουμε σοβαρές ενδείξεις

για να απορρίψουμε την υπόθεση την κανονικότητας με τις εξής εντολές:

```
> par(mfrow=c(1,2))
> hist(Data[,1],xlab="Damage Value",ylab="Frequency",main="Histogram of Damage")
> qqnorm(Data[,1],main="Normal Q-Q Plot of damage",ylab="Sample
  Quantiles",xlab="Theoretical Quantiles")
> qqline(Data[,1])
```



Από τα παραπάνω διαγράμματα βλέπουμε ότι η υπόθεση της κανονικότητας δεν είναι παράλογη.

Για να είμαστε περισσότερο σίγουροι κάνουμε και ένα Shapiro και Wilk normality test

```
> shapiro.test(Data[,1])
```

Shapiro-Wilk normality test

```
data: Data[, 1]
W = 0.97228, p-value = 0.3358
```

Η p-value είναι αρκετά μεγάλη ώστε να μην απορρίψουμε την υπόθεση της Κανονικότητας

Συνεχίζουμε κάνοντας έναν μονόπλευρο έλεγχο με $H_0: \mu_0 = 120$ και εναλλακτική $H_1: \mu_0 > 120$. Η P – τιμή είναι η πιθανότητα δεξιά του T και θα την συγκρίνουμε με το $\alpha=0.05$. Αν η p – value είναι μικρότερη του α , θα απορρίψουμε τη μηδενική υπόθεση. Επίσης, ισοδύναμα, κατασκευάζουμε ένα $(1-\alpha)\% = 95\%$ διάστημα εμπιστοσύνης για τη μέση τιμή του damage δίνοντας στην R την εντολή:

```
> t.test(Data[,1], mu = 120, alternative = "greater")
```

τιμή One Sample t-test

```
data: Data[, 1]
t = 11.274, df = 45, p-value = 5.323e-15
alternative hypothesis: true mean is greater than 120
95 percent confidence interval:
 371.0578      Inf
sample estimates:
mean of x
415.0015
```

ακόμα ότι με πιθανότητα 95% η πραγματική μέση τιμή του damage βρίσκεται μέσα στο διάστημα (371.0578, $+\infty$). Τέλος, ο δειγματικός μέσος βρίσκουμε ότι είναι ίσος με 41.0015. Άρα, ο δημιουργός του παιχνιδιού ότι σε επίπεδο σημαντικότητας 5% είναι σωστός ότι στο συγκεκριμένο στάδιο του παιχνιδιού που διεξάγεται η συγκεκριμένη μάχη, η μέση τιμή της ζημιάς που προκαλεί κάποιος παίκτης είναι μεγαλύτερη από 120 μονάδες.

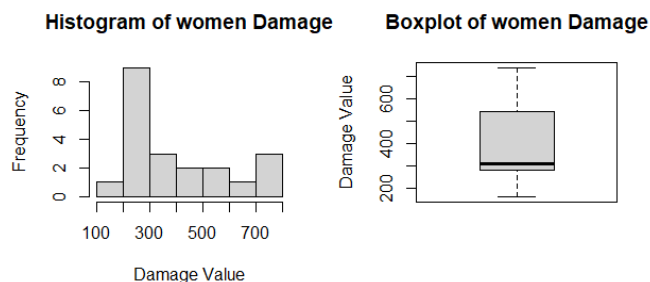
Ένας από τους προγραμματιστές του παιχνιδιού πιστεύει ότι αν ο χαρακτήρας που έχει επιλεγεί είναι γυναίκα, τότε αυτός προκαλεί μεγαλύτερη μέση ζημιά συγκριτικά με τις άλλες δύο επιλογές.

Δημιουργούμε, λοιπόν, την νέα κατηγορική μεταβλητή Woman που αφορά τις θέσεις για τις οποίες η τιμή του sex είναι ίση με “woman” και τις αποθηκεύουμε σε ένα διάνυσμα x. Γράφουμε τον εξής κώδικα:

```
> x <- which(Data$sex=="woman")
> Woman <- c(rep(0,length(Data[,2])))
> Woman[x] <- 1
> Woman <- factor(Woman)
> levels(Woman)
> Data <- transform(Data, woman=Woman)
> Testing_woman <- subset(Data, Woman==1)
> Testing_notWoman <- subset(Data, Woman==0)
```

Κατά την οποία περιγράφουμε το damage αναλόγως με το αν πρόκειται για χαρακτήρα γυναίκας ή όχι. Οπότε χωρίζουμε το δείγμα μας σε δύο subsets ένα που περιέχει τις τιμές για το damage των γυναικών και ένα αυτό των υπόλοιπων χαρακτήρων.

Όσον αφορά τις γυναίκες παίρνουμε τις απαραίτητες πληροφορίες κατασκευάζοντας ένα Ιστόγραμμα, ένα Θηκόγραμμα και για καλύτερη ακρίβεια υπολογίζοντας τις αντίστοιχες αριθμητικές τιμές. Οπότε:



πως η ελάχιστη τιμή είναι λίγο κάτω από 170. Το 25% των τιμών φαίνεται να είναι μεγαλύτερο από περίπου 280, η διάμεσος φαίνεται να είναι λίγο μεγαλύτερη του 300, ενώ το 25% των τιμών φαίνεται να είναι μεταξύ του 500 και του 550. Η μεγαλύτερη τιμή του

Το παραπάνω αποτέλεσμα μας λέει πως η

που παίρνει το στατιστικό ελέγχου μας, T, είναι ίση με 11.274 και πως ακολουθεί κατανομή Student με 45 βαθμούς ελευθερίας. Η p – τιμή είναι πολύ μικρότερη του $\alpha=0.05$, οπότε έχουμε στατιστικές ενδείξεις απόρριψης της μηδενικής υπόθεσης του $\mu_0 = 120$. Ο έλεγχος μας δείχνει

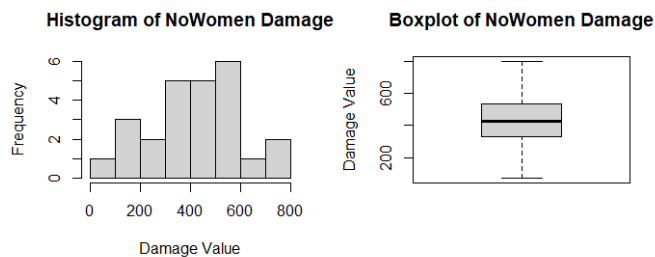
Από το Ιστόγραμμα φαίνεται πως οι περισσότερες γυναίκες έκαναν από 200 έως 300 damage. Η κάπως ανομοιόμορφη κατανομή του damage μας προϋδεάζει για το γεγονός πως η διάμεσος θα είναι πολύ πιο κοντά στο Q1 απ' ό,τι στο Q3, όπως φαίνεται και από το Θηκόγραμμα. Επίσης είναι εμφανές

δείγματος φαίνεται να είναι λίγο μεγαλύτερη του 730. Τέλος, παρατηρούμε πως δεν υπάρχουν έκτροπες τιμές.

```
> fivenum(Testing_woman[,1])
[1] 162.6240 280.6734 311.9488 543.9428 739.9265
> summary(Testing_woman[,1])
Min. 1st Qu. Median Mean 3rd Qu. Max.
162.6 280.7 311.9 402.2 543.9 739.9
> var(Testing_woman[,1])
[1] 32277.92
> sd(Testing_woman[,1])
[1] 179.6606
> IQR <- quantile(Testing_woman[,1],0.75)- quantile(Testing_woman[,1],0.25)
75%
263.2695
```

Όλα τα παραπάνω επιβεβαιώνουν τα συμπεράσματα μας από τα διαγράμματα.

Όσον αφορά τους χαρακτήρες που δεν είναι γυναίκες παίρνουμε τις απαραίτητες πληροφορίες κατασκευάζοντας ένα Ιστόγραμμα, ένα Θηκόγραμμα και για καλύτερη ακρίβεια υπολογίζοντας τις αντίστοιχες αριθμητικές τιμές. Οπότε:



Από το Ιστόγραμμα φαίνεται πως οι περισσότεροι χαρακτήρες με φύλο διάφορο του woman έκαναν από 300 έως 600 damage. Από το Θηκόγραμμα φαίνεται πως η ελάχιστη τιμή είναι κάτω από 100. Το 25% των τιμών φαίνεται να είναι μεγαλύτερο από περίπου 320,

η διάμεσος φαίνεται να είναι λίγο μεγαλύτερη του 400, ενώ το 75% των τιμών φαίνεται να είναι μεταξύ του 500 και του 550. Η μεγαλύτερη τιμή του δείγματος φαίνεται να είναι λίγο μεγαλύτερη του 750. Παρατηρούμε πως δεν υπάρχουν έκτροπες τιμές.

```
> fivenum(Testing_woman[,1])
[1] 74.91456 328.02945 427.83549 531.87218 796.87913
> summary(Testing_woman[,1])
Min. 1st Qu. Median Mean 3rd Qu. Max.
74.91 328.03 427.84 425.75 531.87 796.88
> var(Testing_woman[,1])
[1] 31889.49
> sd(Testing_woman[,1])
[1] 178.5763
> IQR <- quantile(Testing_woman[,1],0.75)- quantile(Testing_woman[,1],0.25)
75%
203.8427
```

Όλα τα παραπάνω επιβεβαιώνουν τα συμπεράσματα μας από τα διαγράμματα.

Στην συνέχεια, κατασκευάζουμε μια νέα κατηγορική μεταβλητή με το όνομα **elf_cat** με δύο κατηγορίες (0: όχι ξωτικό και 1: ξωτικό) και μια νέα κατηγορική μεταβλητή με το όνομα **damage_cat** η οποία παίρνει την “τιμή” **low** όταν το **damage** είναι < 400 και την “τιμή” **high**

όταν το **damage** είναι 400. Θέλουμε να ελέγξουμε σε ε.σ. 5%, κατά πόσο οι εν λόγω τ.μ. είναι ανεξάρτητες.

Γράφουμε τον παρακάτω κώδικα

```
> x <- which(Data$faction=="elf")
> elf_cat <- rep(0, length(Data[,3]))
> elf_cat[x] <- 1
> elf_cat <- factor(elf_cat)

> x <- which(Data[,1]<400)
> damage_cat <- rep("high", length(Data[,1]))
> damage_cat[x] <- "low"
> damage_cat <- factor(damage_cat)
```

```
> table(elf_cat,damage_cat)
```

	damage_cat	
elf_cat	high	low
0	16	18
1	6	6

Δίπλα φαίνεται ο πίνακας συνάφειας. Ως μηδενική υπόθεση, H_0 , παίρνουμε ότι οι elf_cat και damage_cat είναι ανεξάρτητες, ενώ ως εναλλακτική υπόθεση, H_1 , έχουμε ότι οι elf_cat και damage_cat δεν είναι ανεξάρτητες. Αφού όλες οι τιμές του πίνακα συνάφειας είναι μεγαλύτερες του 5, μπορούμε να κάνουμε ένα X^2 τεστ ανεξαρτησίας χρησιμοποιώντας την παρακάτω εντολή:

```
> chisq.test(table(elf_cat,damage_cat))
```

Pearson's Chi-squared test with Yates' continuity correction

data: table(elf_cat, damage_cat)

X-squared = 2.6345e-31, df = 1, p-value = 1
απαντάμε ότι σε ε.σ. 5% δεν υπάρχει καμία σοβαρή ένδειξη που να υποστηρίζει ότι το damage_cat και το elf_cat δεν είναι ανεξάρτητα.

Ο έλεγχος μας δίνει μία πάρα πολύ υψηλή p – value, επομένως δεν έχουμε στατιστικά σημαντικές ενδείξεις ώστε να απορρίψουμε τη μηδενική υπόθεση της ανεξαρτησίας. Άρα,

απαντάμε ότι σε ε.σ. 5% δεν υπάρχει καμία