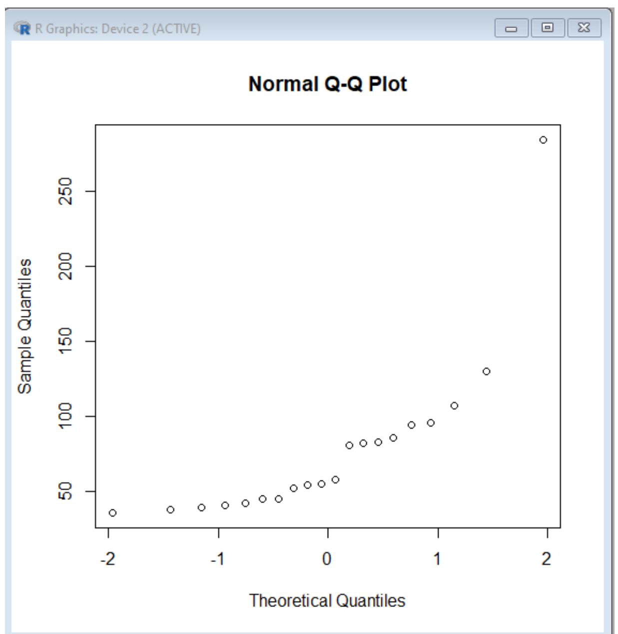
ΠΑΝΤΕΛΟΠΟΥΛΟΣ ΣΤΥΛΙΑΝΟΣ 3180142

ΑΠΟΣΤΟΛΟΠΟΥΛΟΣ ΘΕΜΙΣΤΟΚΛΗΣ 3180013

ΑΣΚΗΣΗ 1

Τα δεδομένα είναι κατάλληλα για να εξάγουμε συμπέρασμα καθώς προέρχονται από τυχαία δειγματοληψία (SRS), είναι σε πλήθος περισσότερα από 15, και ατυπικά σημεία θεωρούμε ότι δεν υπάρχουν καθώς μπορεί να υπάρχουν περισσοτέρες τιμές κοντά στο 284 και να μην έγινε λήψη αυτών. Αυτό μπορούμε να το διακρύνουμε και στο Normal quantile plot καθώς όλα να σημεία, εκτός από το 284, είναι πάνω σε μια ευθεία.



Το 95% διάστημα εμπιστοσύνης για τη μέση τιμή του χρόνου είναι: [51.41365, 103.38635].

Υπολογίστηκε με τη χρήση της γλώσσας R όπου σε ένα διάνυσμα εισάγαμε τα δεδομένα υπολογίσαμε το μήκος του διανύσματος n, βρήκαμε το t<- -qt(0.025, df = n-1) και τέλος με τη χρήση του τύπου για τον υπολογισμό του διαστήματος εμπιστοσύνης

(mean(x) +c(-1,1)*t*sd(x)/sqrt(n)) πήραμε το αποτελεσμά μας.

ΑΣΚΗΣΗ 2

a) α) Το λάθος βρίσκεται στον τύπο της τυπικής απόκλισης του δειγματικού μέσου καθώς στον παρανομαστή θα έπρεπε να είχε ρίζα 20.

b) Β)Το λάθος είναι ότι στη μηδενική υπόθεση υποθέτουμε ότι η μέση τιμή του δείγματος είναι 10 και όχι η μέση τιμή του πλυθησμού που εξετάζουμε.

c) Γ)Το λάθος βρίσκεται στην απόρριψη της μηδενικής υπόθεσης καθώς δεν ξέρουμε αν ισχύει η όχι ο κανόνας της απόφασης (p-value <= a) έτσι ώστε να απορριφθεί η υπόθεση μας.

d) Δ)Το λάθος βρίσκεται στην απόριψη της μηδενικής υπόθεσης καθώς δεν ξέρουμε το επίπεδο σημαντικότητας (α) και αν το p-value είναι μεγαλύτερο από αυτό για να απορρίψουμε την αρχική υπόθεση

ΑΣΚΗΣΗ 3

Βρίσκουμε ότι για z = 1.34, Φ(z)= 0.9099 και έχουμε:

a) Α) 1 – Φ(z)= 0.09122

b) Β) Φ(z)= 0.9099

c) Γ) 2 Φ(-|z|)= 0.18024

ΑΣΚΗΣΗ 4

a) Το επίπεδο σημαντικότητα είναι α= 100% - 95% = 5%.

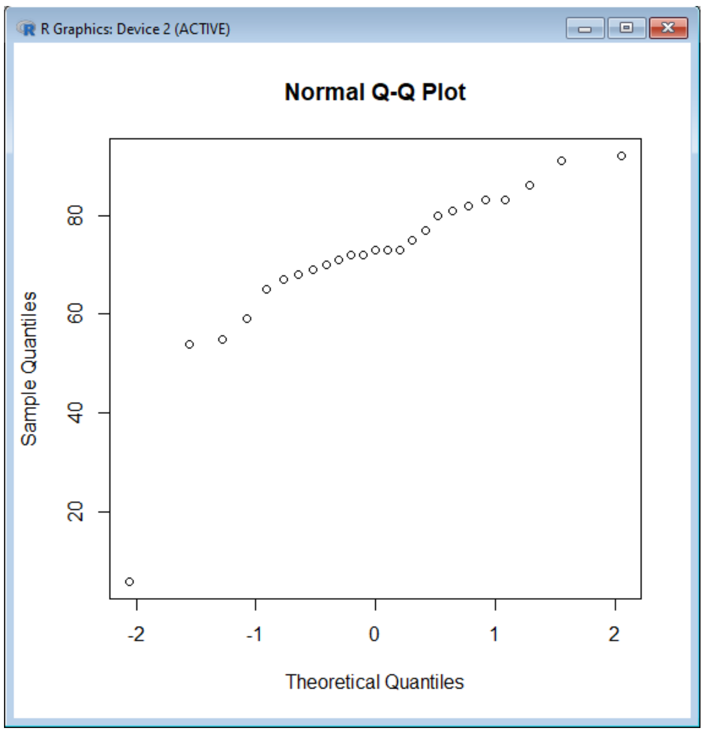
   Επίσης p-value = 4%, συνεπώς ισχύει ότι p-value < a, άρα από τον κανόνα της απόφασης (p-value <= a) απορρίπτεται η μηδενική υπόθεση και δεν ξέρουμε αν το 30 ανήκει στο 95% διάστημα εμπιστοσύνης.

b) Το επίπεδο σημαντικότητας είναι α= 100% - 90% = 10%.

   Άρα πάλι ισχύει ότι p-value < a και μπορούμε από τον κανόνα της απόφασης να απορρίψουμε την μηδενική υπόθεση και κατά συνέπεια να μη ξέρουμε αν το 30 ανήκει στο 90% διάστημα εμπιστοσύνης.

<u>ΑΣΚΗΣΗ 5</u>

Τα δεδομένα μας προέρχονται από τυχαία δειγματοληψία, είναι περισσότερα από 15 αλλά υπάρχουν ατυπικά σημεία (ένα άτομο έχε δηλωθεί πως έχει βάρος 6 κάτι το οποίο είναι αδύνατο για έναν ενήλικα) συνεπώς αν αφαιρέσουμε το ατυπικό σημείο, μπορούμε να τα χρησιμοποιήσουμε για να εξάγουμε συμπεράσματα. Αυτό μπορούμε να το διακρύνουμε και στο Normal quantile plot στο οποίο αν αφαιρέσουμε το ατυπικό σημείο τα δεδομένα μας βρίσκονται πάνω σε μια ευθεία.



a) Το 95% διάστημα εμπιστοσύνης για το μέσο βάρος των ενηλίκων στην Αθήνα είναι: [69.57826, 78.00507]

b) Το 80% διάστημα εμπιστοσύνης για τη διαφορά του μέσου βάρους των αντρών και των γυναικών είναι: [5.948055, 15.436561].

c) Δημιουργούμε δύο πίνακες έναν με το βάρος των καπνιστών και έναν με το βάρος των μη καπνιστών και με μηδενική υπόθεση ότι το μέσο βάρος των καπνιστών ισούται με το μέσο βάρος των μη καπνιστών και εναλλακτική υπόθεση ότι δεν ισούται έχω

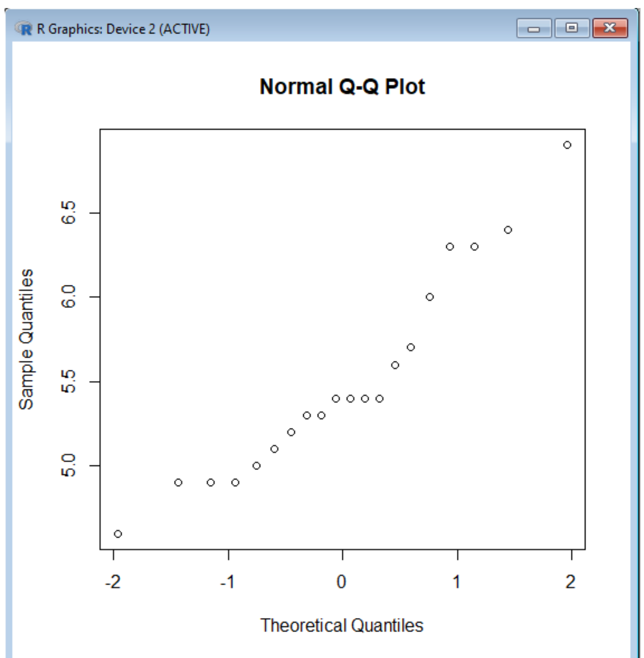
H_0: $\mu_\kappa = \mu_{\mu\kappa}$

H_α: μ_κ != $\mu_{\mu\kappa}$

Όπου μ_κ μέσο βάρος καπνιστών και $\mu_{\mu\kappa}$ μέσο βάρος μη καπνιστών. Κάνοντας t-test με τις συγκεκριμένες υποθέσεις παίρνουμε p-value = 0.2228 αρκετά μεγάλο συνεπώς απορρίπτουμε τη μηδενική υπόθεση και συμπαιρένουμε ότι το βάρος έχει σχέση με το κάπνισμα.

ΑΣΚΗΣΗ 6

a) Τα δεδομένα προέρχονται από τυχαία δειγματοληψία, δεν υπάρχουν ατυπικά σημεία και το πλήθος των δεδομένων είναι μεγαλύτερο από 15. Επίσης βλέπουμε ότι τα δεδομένα είναι κατάλληλα για τις μέθοδους συμπερασματολογίας και από το Nomral quantile plot καθώς όλα τα σημεία σχηματίζουν μια ευθεία.



b) Μέση τιμή: 5.5

Τυπική απόκλιση: 0.6008766

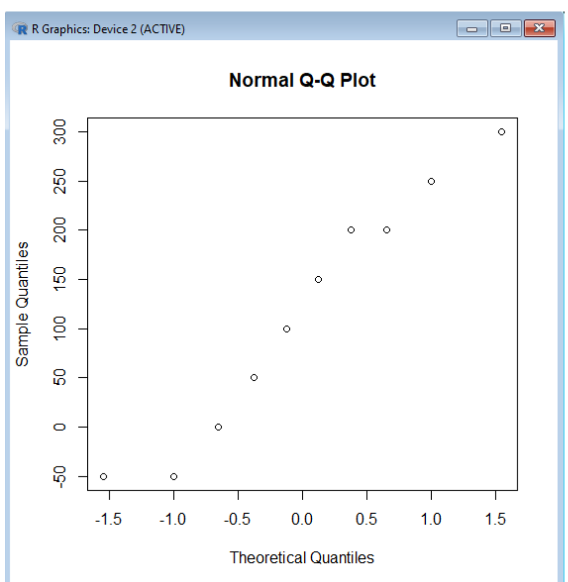
c) Το 95% διάστημα εμπιστοσύνης για τη μέση τιμή κατανάλωσης καυσίμου για τα αυτοκίνητα είναι: [5.218781, 5.781219].

<u>ΑΣΚΗΣΗ 7</u>

Δημιουργούμε έναν νέο πίνακα ο οποίος θα περιέχει τη διαφορά των δύο τιμών(ΣΥΝΕΡΓΕΙΟ – ΕΜΠΕΙΡΟΓΝΩΜΟΝΑΣ).

ΑΥΤΟΚΙΝΗΤΟ	1	2	3	4	5	6	7	8	9	10
ΔΙΑΦΟΡΑ ΤΙΜΗΣ	100	50	-50	0	-50	200	250	200	150	300

Τα δείγματα είναι τυχαία και όπως φαίνεται και από το Normal quantile plot δεν υπάρχουν ατυπικά σημεία συνεπώς μπορούμε να χρησιμοποιήσουμε τα δεδομένα αυτά για εξαγωγή συμπερασμάτων.



Για να δούμε αν το συνεργείο υπερκοστολογεί τις επισκευές θα χρησιμοποιήσουμε τον μέσο όρο της διαφοράς των τιμών. Θέτουμε μηδενική υπόθεση ότι ο μέσος όρος ισούται με μηδέν και έχουμε:

H_0: μ=0

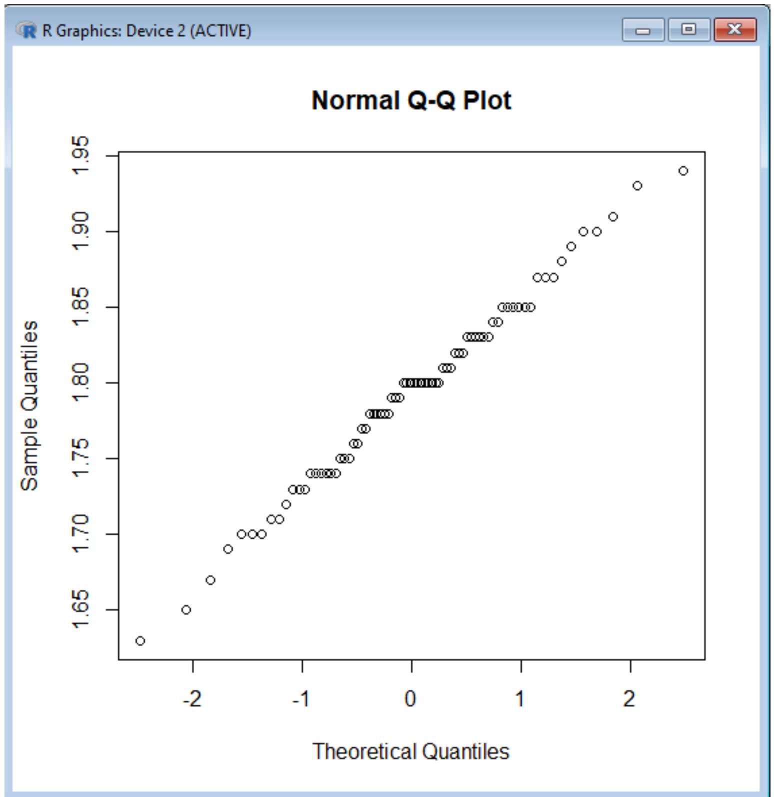
$H_α$: μ > 0

όπου μ είναι ο μέσος όρος της διαφοράς.

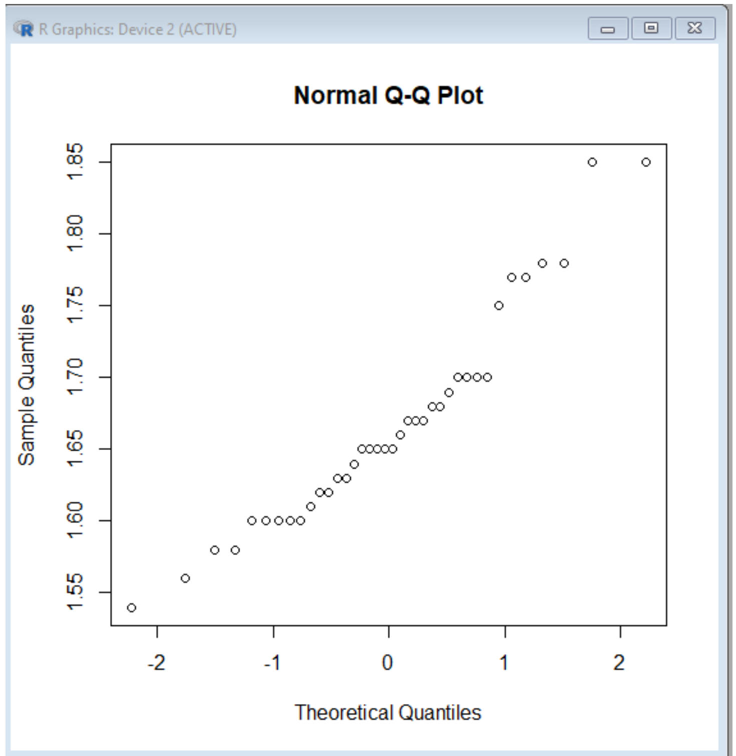
Θα βρούμε το p-value και αν είναι πολύ μικρό τότε θα απορρίψουμε την μηδενική υπόθεση με βάση των κανόνα της απόφασης (p-value<= a). Πραγματοποιώντας ένα t-test() βλέπουμε ότι p-value = 0.008611, πολύ μικρό ακόμα και για επίπεδο εμπιστοσύνης α=1%. Συνεπώς απορρίπτουμε την μηδενική υπόθεση.

<u>ΑΣΚΗΣΗ 8</u>

a) Ελέγχοντας τα δεδομένα για τον ανδρικό πλυθησμό του τμήματος παρατηρούμε πως υπάρχει ένα ατυπικό σημείο το 176.00 καθώς έγινε τυπογραφικό λάθος αλλάζοντας το 176.00 σε 1.76 παρατηρούμε ότι τα δεδομένα μας είναι κατάλληλα για εξαγωγή συμπερασμάτων κάτι το οποίο παρατηρούμε και από το Normal quantile plot.

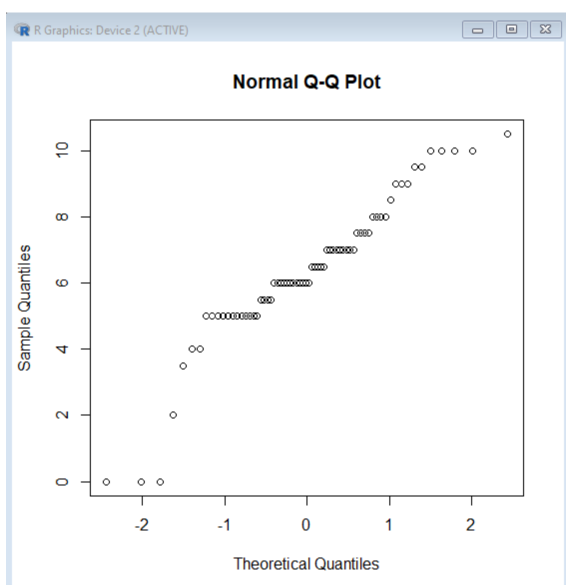


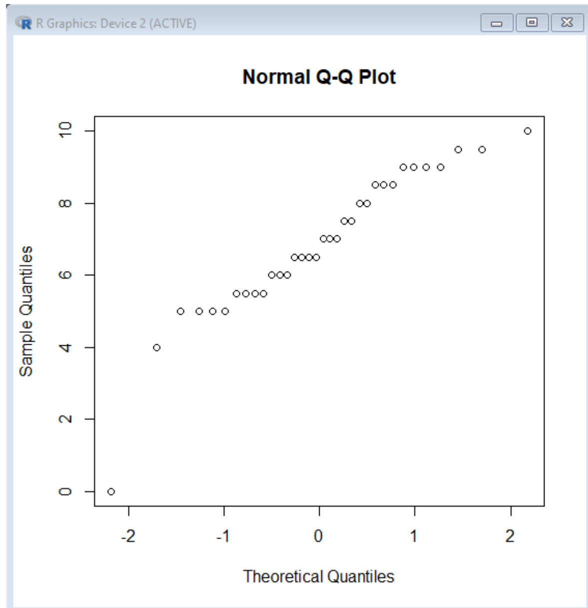
Τα δεδομένα για τον γυναικείο πλυθησμό δεν χρειάζονται κάποια επεξεργασία καθώς έχουν δοθεί ορθά και είναι κατάλληλα για εξαγωγή συμπερασμάτων.



Το 95% διάστημα εμπιστύνης για τη διαφορά του μέσους ύψους μεταξύ αντρών και γυναικών στο τμήμα πληροφορικής είναι: [0.09846512, 0.15469278].

b) Από τα Normal quantile plots (το πρώτο είναι για τους άντρες το δεύτερο για τις γυναίκες) βλέπουμε ότι τα δεδομένα μας μπορούν να χρησιμοποιηθούν για εξαγωγή συμπερασμάτων.





Θέτω μηδενική υπόθεση ότι ο μέσος βαθμός των αντρών είναι ίσος από τον μέσο όρο των γυναικών και έχω:

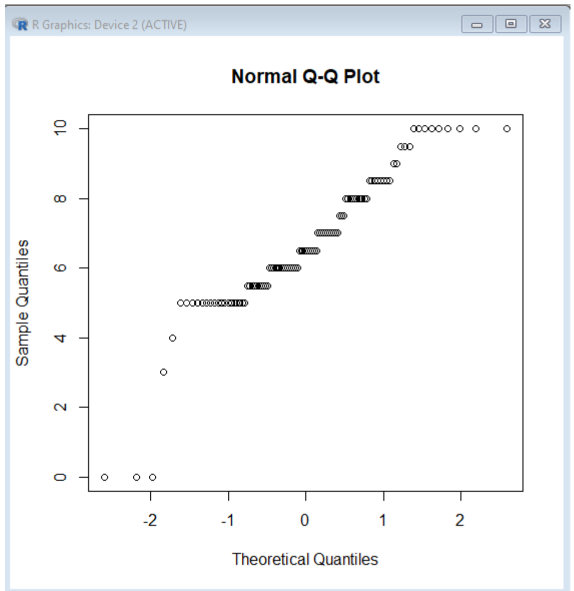
H_o: $\mu_A = \mu_\Gamma$

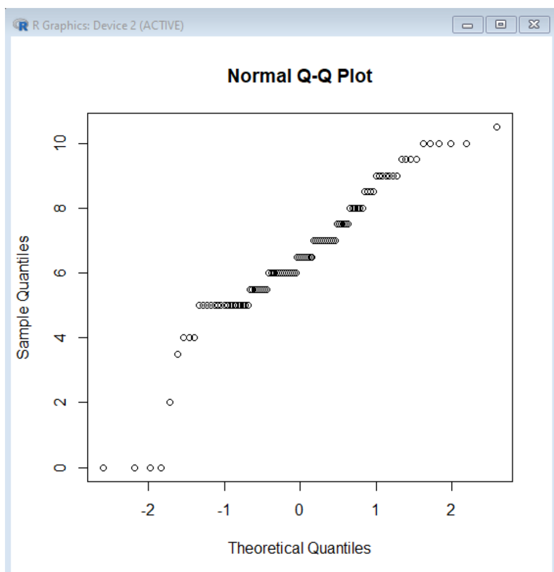
H_a: $\mu_A > \mu_\Gamma$

Θα υπολογίσουμε το p-value και συγκρίνοντας το με το επίπεδο σημαντικότητας α = 5% θα συμπεράνουμε αν πρέπει να απορρίψουμε την μηδενική υπόθεση η όχι.

Μέσα από το t-test που πραγματοποιήσαμε βρήκαμε ότι p-value = 0.1202 και έχουμε ότι p-value > α συνεπώς από τον κανόνα της απόφασης μπορούμε να αποδεχτούμε την μηδενική υπόθεση.

c) Παρατηρώντας τα δεδομένα καθώς και τα Normal quantile plots (το πρώτο αφορά τους βαθμούς στα μαθηματικά1 και το δεύτερο τους βαθμούς στις πιθανότητες) τους βλέπουμε πως είναι κατάλληλα για εξαγωγή συμπερασμάτων.





Θέτοντας μηδενική υπόθεση ότι οι δύο μέσοι βαθμοί είναι ίσοι και εναλλακτική οτί είναι διαφορετικοί

H_0: $\mu_\mu = \mu_\pi$

H_α: μ_μ != μ_π

με μ_μ να είναι ο μέσος βαθμός στα Μαθηματικά 1 και μ_π να είναι ο μέσος βαθμός στις πιθανότητες.

Κάνοντας t-test με αυτές τις υποθέσεις παίρνουμε σαν αποτέλεσμα p-value = 0.5445 το οποίο είναι πολύ μεγάλο κάτι που σημαίνει ότι η μηδενική υπόθεση ισχύει.