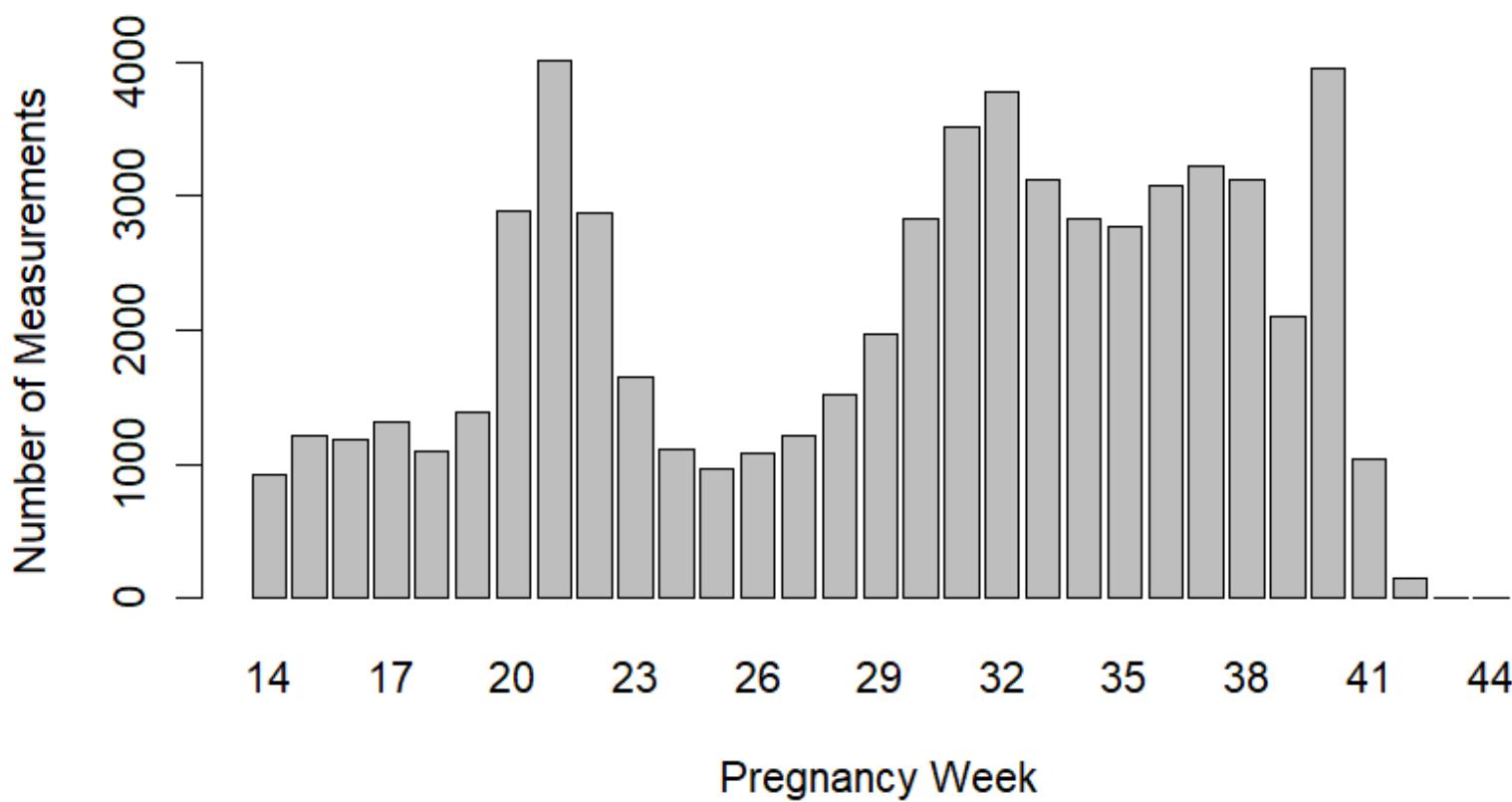


יצירת עקומות גדילה לעוברים

קובץ AC3
מטלה 1

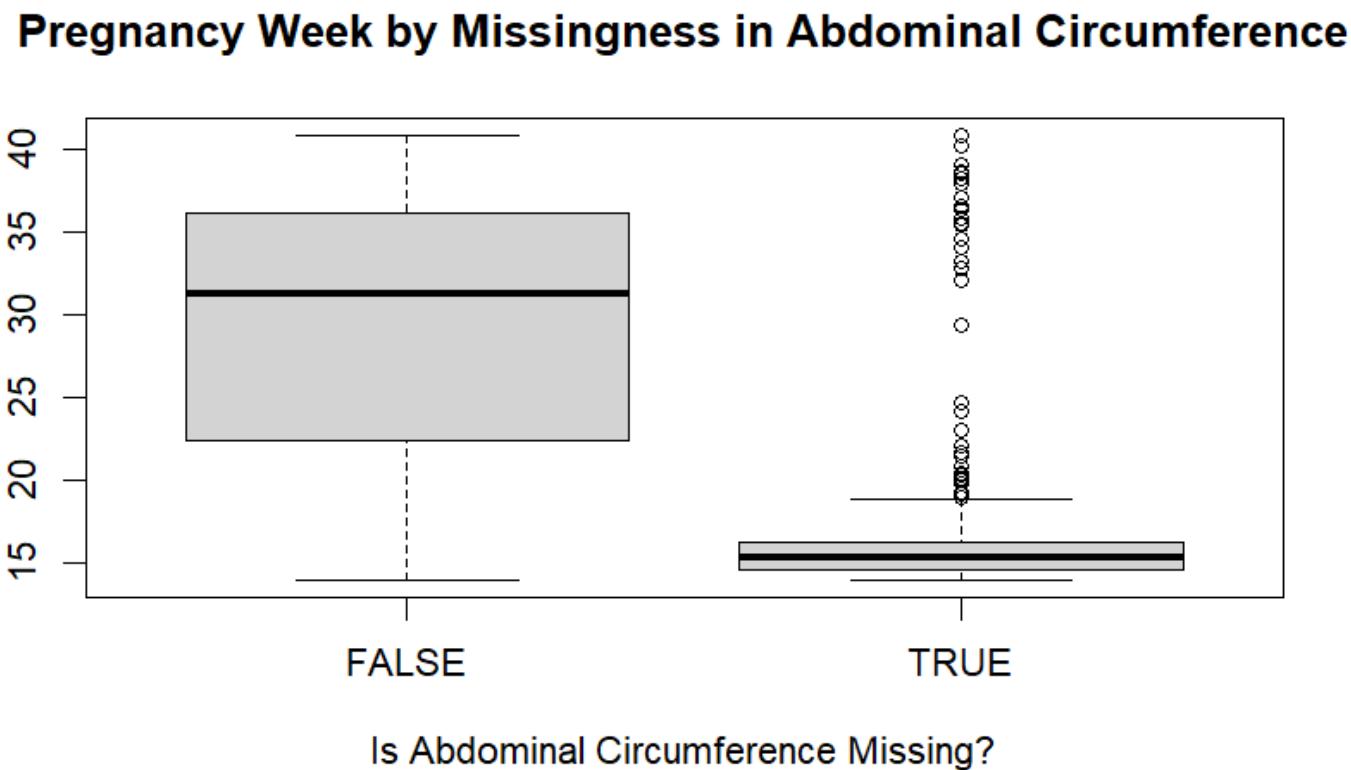
הגבלת הנתונים ומספר התצפיות

Number of Measurements per Pregnancy Week



- ❖ בתרשים ניתן לראות את מספר המדידות עבור כל שבוע בהריון.
- ❖ הגבלנו את הנתונים עד לשבוע 40 (כולל), כדי שהתבקשנו לעשות, והורדנו את המדידות שהתקיימו מהשבוע ה-41 (כולל) ואילך - סך הכל הסרנו 1,208 תצפיות.
- ❖ בהתקדמות הנתונים עד השבוע ה-40 (כולל), אנחנו מסתכלים על 15,451 נשים ועל 24,138 הריונות.

מדידות חסרות



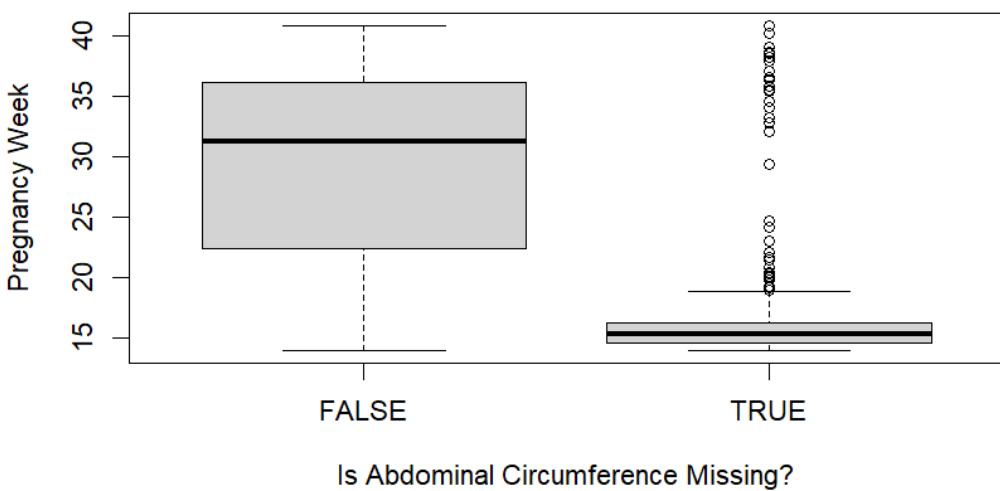
- ❖ בתרשים מוצגים החסרים במדידות היקף הבطن של העוברים לפי שבוע ההריון
- ❖ לפניו שנזורך ערכים חסרים, נבדוק האם יש קשר בין המשתנה החסר (AC) לבין שבוע ההריון?
- ❖ מהבוקס-פלוט אפשר לראות שיש קשר משמעותי בין השניים:
 - כאשר הדירוג cir_ab לא חסר — (FALSE) התצפויות משתרעות על פני כל טווח ההריון.
 - כאשר הדירוג cir_ab חסר — (TRUE) רוב התצפויות הן בהריוןות מוקדמים מאוד.
- ❖ ננסה להסביר, מהם בדיק השבועות של ההריון המוקדמים? כי הבוקס-פלוט נותן מידע מוגבל בעמודות TRUE

מדידות חסרות

> `bp$stats`

	[,1]	[,2]
[1,]	14.00000	14.00000
[2,]	22.42857	14.57143
[3,]	31.28571	15.28571
[4,]	36.14286	16.28571
[5,]	40.85714	18.85714

Pregnancy Week by Missingness in Abdominal Circumference



❖ עמודה 1 בטבלה מייצגת את FALSE ועמודה 2 את TRUE, כאשר: שורה 1 מייצגת את הערך המינימלי; שורה 2 את הרביעון הראשון; שורה 3 את החציון; שורה 4 את הרביעון השלישי; ושורה 5 את הגבול העליון לפני הערכים החרייגים.

❖ ניתן לראות כי המדידה הנמוכה ביותר של שבוע הירון בתצפיות החסרות היא שבוע 14.

❖ חצי מהתצפיות החסרות הן עד שבוע 15.3
❖ השפיף העליון הוא 18.86, ככלומר הגבולה ביותר לפני שטחחים החרייגים הוא בערך שבוע 19.

❖ השפיף העליון הוא הערך הגדול ביותר שעדיין נחשב "לא חרייג", ככלומר כל ערך שהוא קטן או שווה ל-: $IQR * 1.5 + Q3 - Q1 = Q3 - Q1$ אשר $IQR = Q3 - Q1$

❖ כל ערך שמעל הגבול הזה נחשב חרייג (outlier) ומציין בתור עיגול.

❖ מאחר שהחסרים ב- AC מתרכזים כמעט אך ורק בשבועות 14–18, והמדידות הופכות עקביות ורציפות רק החל משבוע 19, בחרנו לנתח את הנתונים החל משבוע זה.

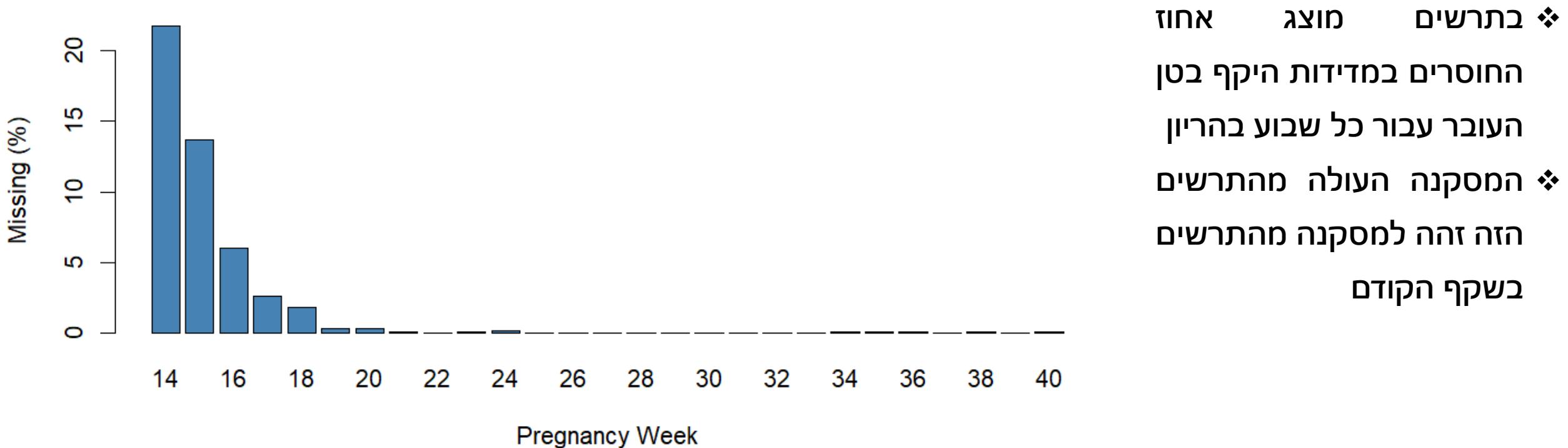
התמודדות עם מדידות חסרות

Week	Total	NonMissing	Missing	MissingPercent
14	14	929	727	202
15	15	1207	1042	165
16	16	1179	1108	71
17	17	1320	1285	35
18	18	1103	1083	20
19	19	1383	1378	5
20	20	2894	2885	9
21	21	4016	4012	4
22	22	2872	2871	1
23	23	1654	1653	1
24	24	1111	1109	2
25	25	972	972	0
26	26	1084	1084	0
27	27	1220	1220	0
28	28	1516	1516	0
29	29	1973	1972	1
30	30	2835	2835	0
31	31	3512	3512	0
32	32	3782	3780	2
33	33	3130	3129	1
34	34	2826	2824	2
35	35	2781	2778	3
36	36	3078	3076	2
37	37	3233	3232	1
38	38	3119	3115	4
39	39	2100	2099	1
40	40	3962	3958	4

- ❖ בטבלה מוצגים נתוני המדידות הקיימות והחסרות של היקף בתן העובר עבור כל שבוע, מהשבוע ה-14 ועד השבוע ה-40
- ❖ בנוסף, ב العمودה הימנית, מוצג אחוז המדידות החסרות עבור שבוע ההריאון המתאים
- ❖ ניתן לראות כי בשבועות 18-14 אחוז המדידות החסרות הוא גבוה, וחלל שבוע 19 אחוז המדידות החסרות נמוך יותר
- ❖ לכן החלטנו להתחילה את הניתוח שבוע 19 (כולל)

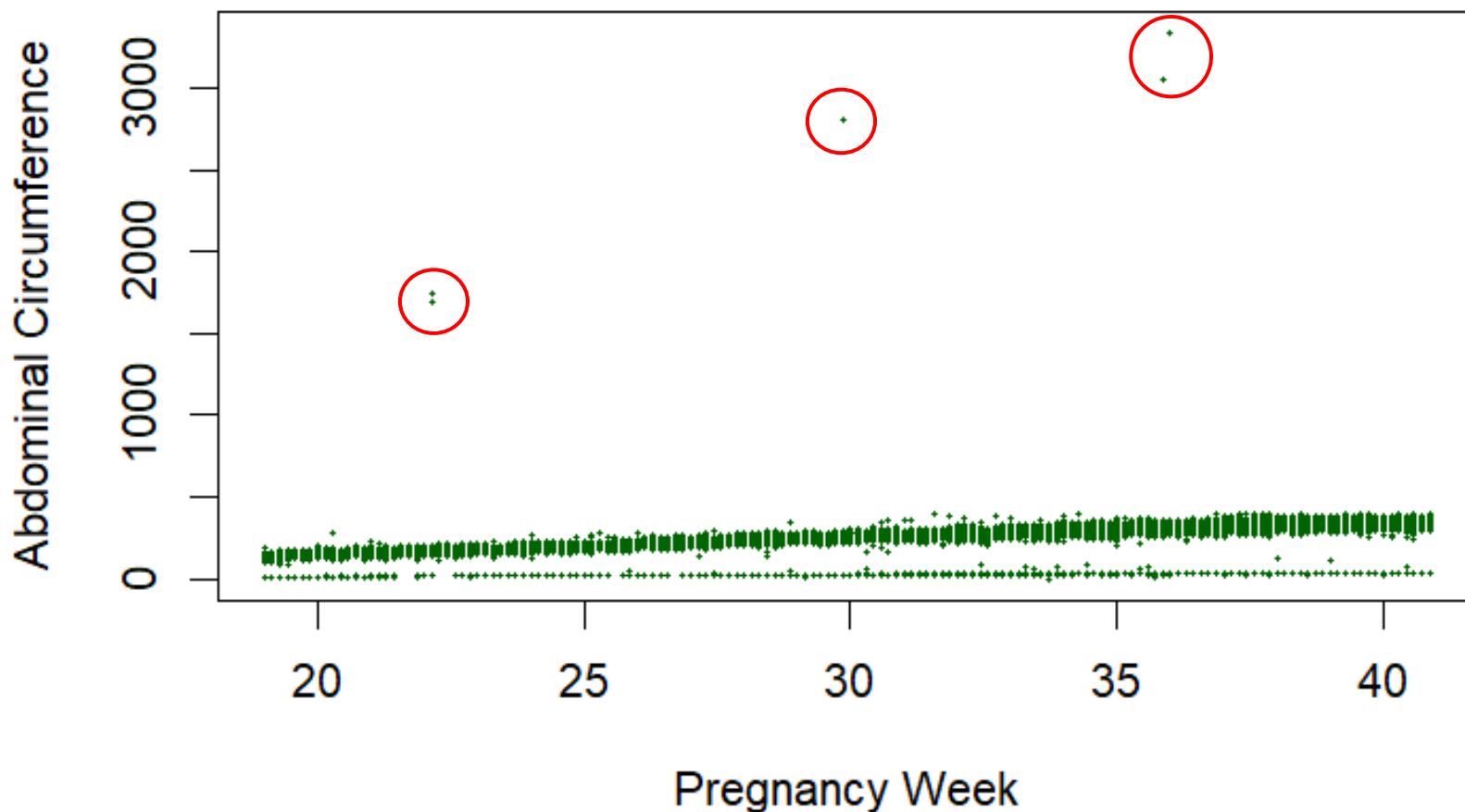
התמודדות עם מדידות חסרות

Percentage of Missing Abdominal Circumference Values by Pregnancy Week



מדידות חריגות ייחידות מידת

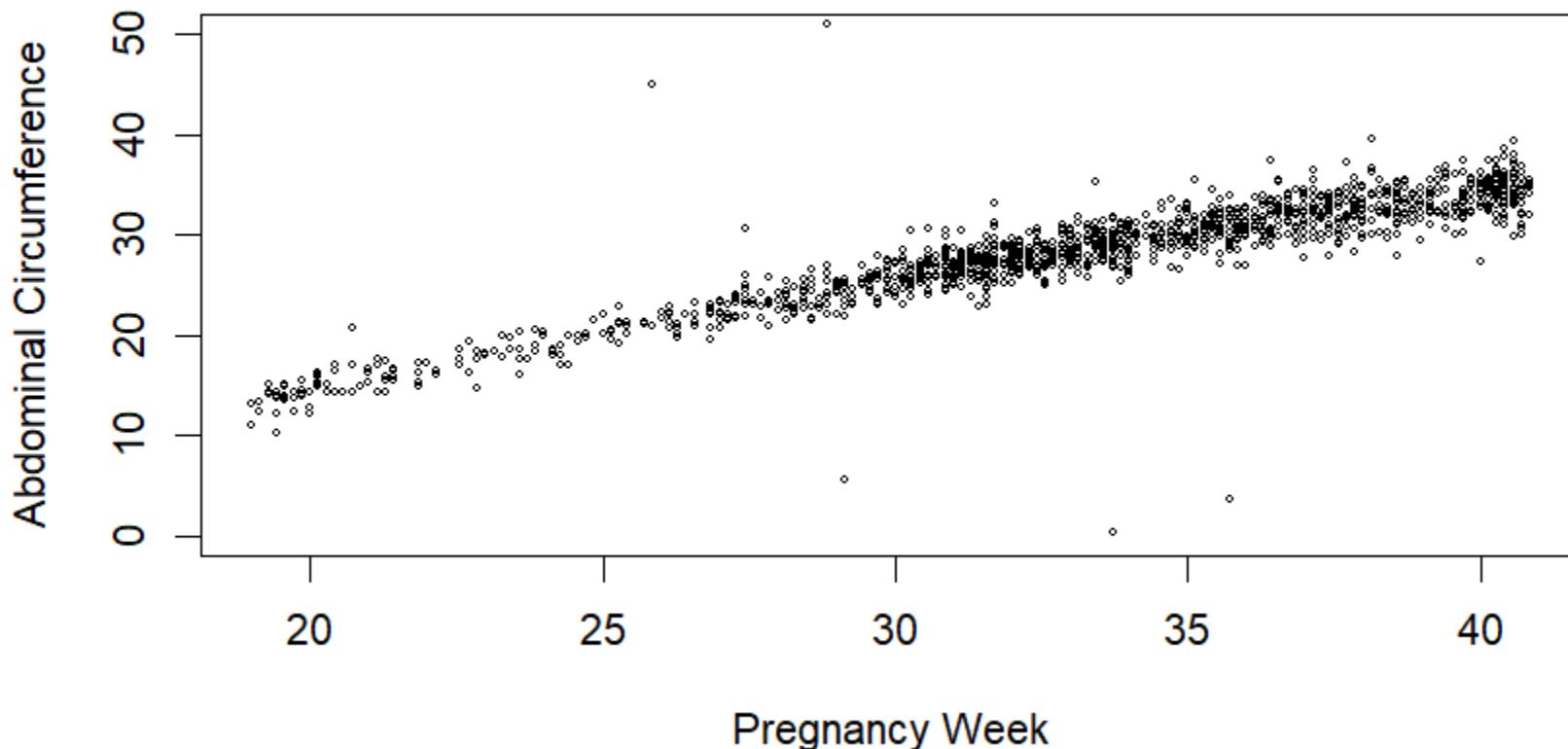
Abdominal Circumference vs. Pregnancy Week (≥ 19)



- ❖ בתרשים מוצגות מדידות היקף הבطن של העובר נגד שכעות ההריון.
- ❖ יש מספר קטן של תצפיות חריגות (outliers), שמסוכנות באדום, בהן מדידות היקף בطن העובר הוא מעל 1000. החלנו להסiran.
- ❖ לא הוספנו ייחידות מידת גלגל שה"עננים" השונים שמוסגים בתרשימים מדגימים את הפער הקיים ביחסות המידה.

מדידות חריגות ייחידות מידת

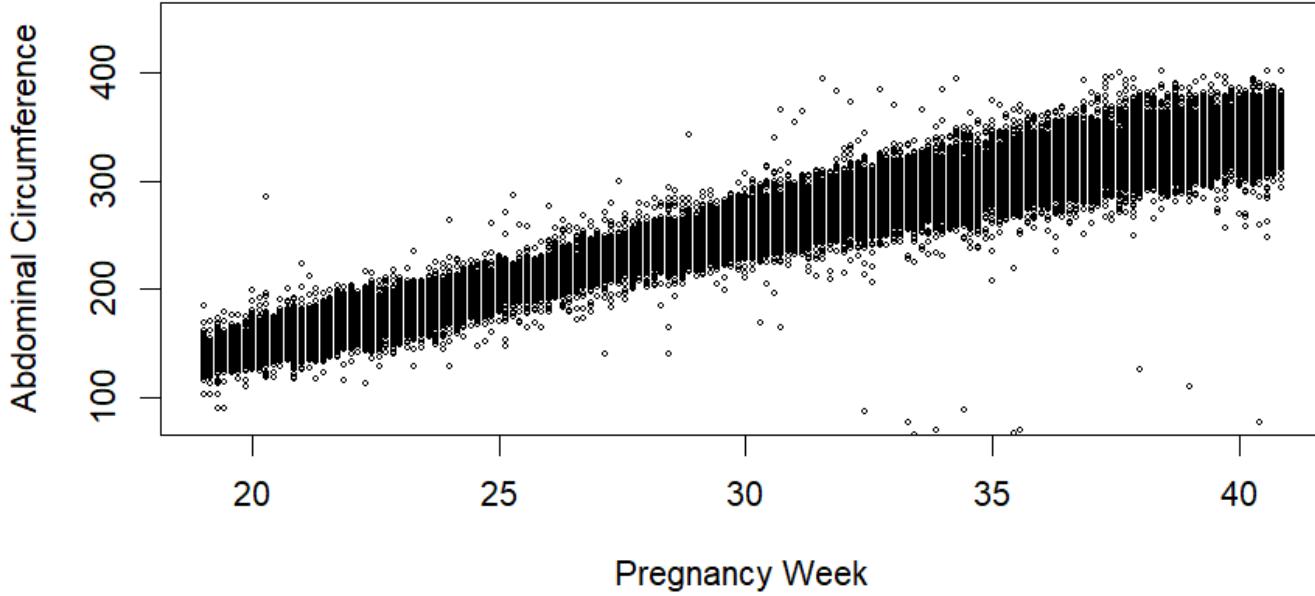
Abdominal Circumference vs Pregnancy Week (Zoomed In: 0–50)



- ❖ בתרשים מוצגות מדידות היקף בטן העובר לאורך שבועות ההריון.
- ❖ בגלל התוצאות החריגות שモוצגות בתרשימים בשקף הקודם, לא יוכלו לראות מה ייחידות המידה.
- ❖ עשינו זום-אין כדי לראות את הענוקתן מהתרשימים הקודם בבחירה.

מדידות חריגות ויחידות מידע

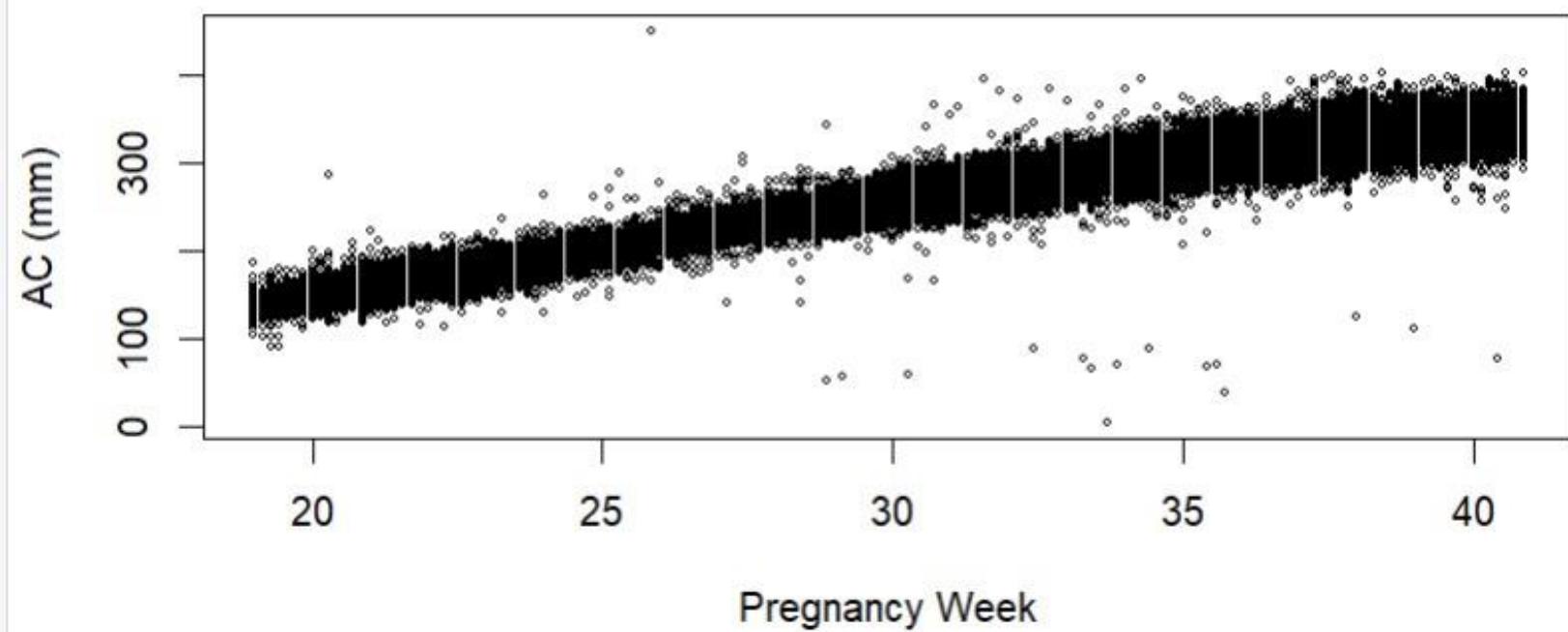
Abdominal Circumference vs Pregnancy Week (Zoomed In: 80–450)



- ❖ בתרשימים מוצגות מדידות היקף בטן העובר לאורך שבועות ההריון.
- ❖ בגלל התכפיות החריגות שモוצגות בתרשימים בשקף 7, לא יכוליםנו לראות מה ייחידות המידע.
- ❖ עשינו זום-אין כדי לראות את הענן הגדל יותר מהתרשימים הקודם בבירור.
- ❖ אחרי שבחנו את התרשימים בשקף זהה ובשקף הקודם, הסכנו כי בחלוקת מהמדידות השתמשו בס"מ ובחלוקת מהמדידות בmm".
- ❖ כדי לגשר על הפער, המרנו את כלל המדידות לmm".

התמודדות עם מדידות חריגות

Abdominal Circumference (Standardized to mm)



- ❖ בתרשים מוצגות מדידות היקף בטן העובר (מ"מ) לאורך שבועות ההריון.
- ❖ בדקנו כמה ערכים חריגים נשארו לנו, ככלمر בכמה מהמדידות היקף בטן העובר קטן מ-100, או לחילופין גדול מ-400.
- ❖ סה"כ קיבלנו ש-19 מהתצפיות הן חריגות – מספר זנich מתוך 55,005 תצפיות סך הכל.

בדיקות הנוסחה של הדלוק למדידת היקף בטן העובר

$$AC = -13.3 + 1.61 \times GA - 0.00998 \times GA^2$$

Input Unit: w (week)

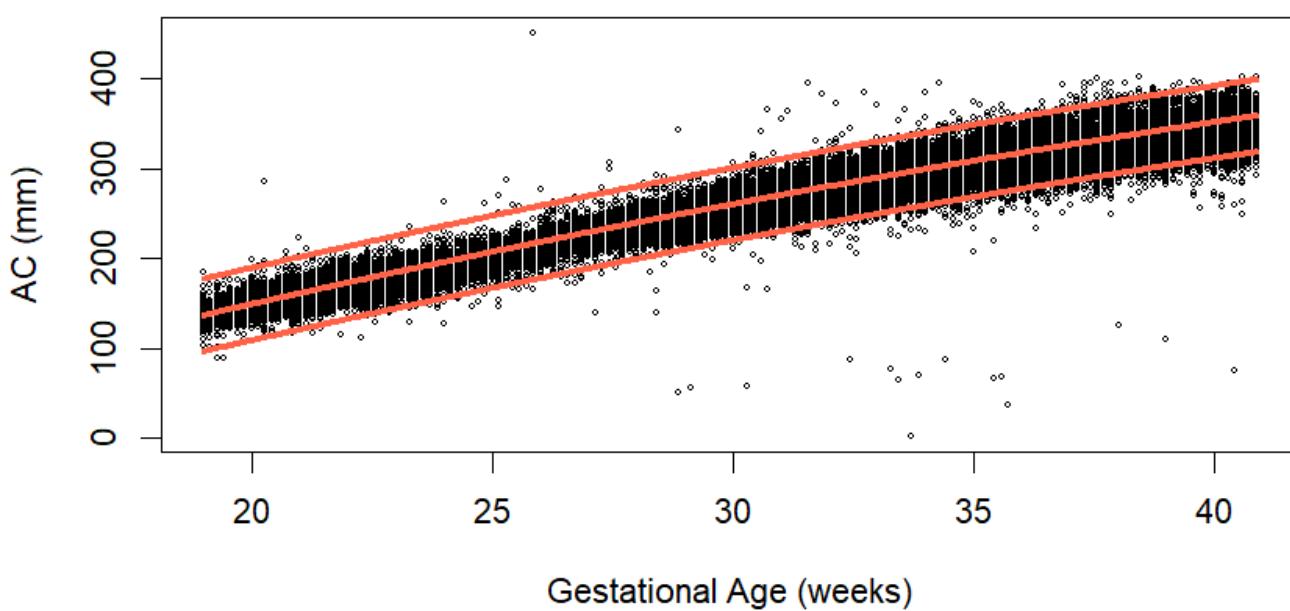
Output Unit: cm

Min Range: 12.0 weeks

Max Range: 40.0 weeks

Standard Deviation (\pm): 1SD = 1.34 cm

Observed AC (mm) vs Hadlock $\pm 3SD$



- ❖ בתרשים העליון מוצגת נוסחת הדלוק לחישוב הערך הצפוי של היקף בטן לפי גיל ההריון.
- ❖ ייחדות המידה של הקלט הן שבועות, ייחדות המידה של הפלט הן ס"מ. סטיית התקן היא 1.34 ס"מ.
- ❖ אנחנו עבדנו בס"מ ולכן בנוסחה שעשינו הכפלנו ב- $.10$.
- ❖ כעת נתיחס לקוים האדומים שמופיעים בתרשים הפיזור שלנו: הקו האמצעי הוא הערך הצפוי לפי הנוסחה, הקווים החיצוניים מייצגים את חום של $+/- 3$ סטיות תקן.
- ❖ ניתן לראות שנוסחת הדלוק ככל הנראה לא מתאימה לתוצאות שלנו – נבחן זאת לעומק.

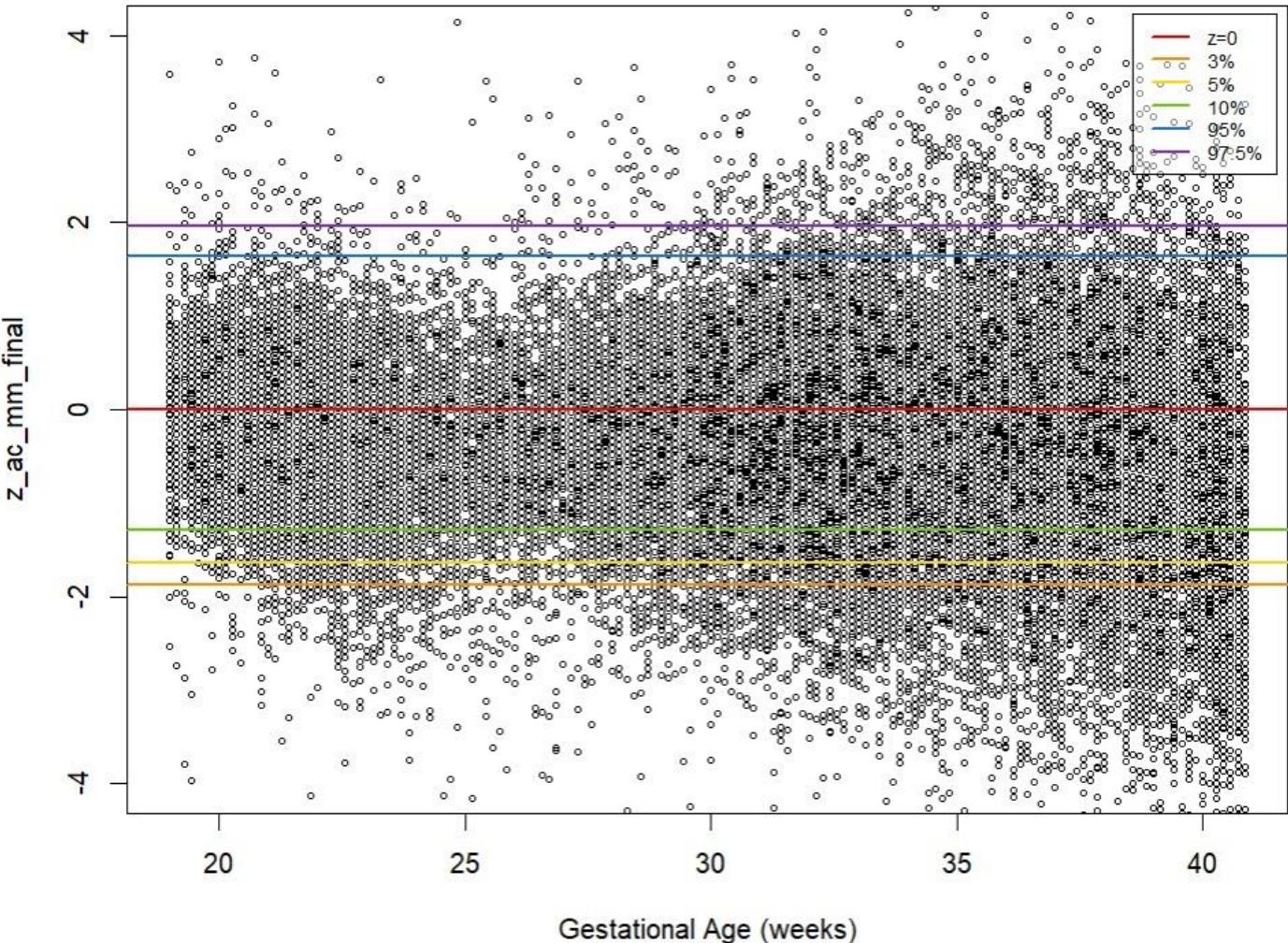
בדיקות התאמה להתפלגות נורמלית סכיב מודל הדלק

	Tail	Observed	Expected	Lower_CI	Upper_CI
1	Below 3%	0.08139260	0.030	0.02854529	0.03145471
2	Below 5%	0.11324425	0.050	0.04814144	0.05185856
3	Below 10%	0.18251068	0.100	0.09744171	0.10255829
4	Above 95%	0.02757931	0.050	0.04814144	0.05185856
5	Above 97.5%	0.01641669	0.025	0.02366862	0.02633138

- ❖ ביצענו ניתוח של שיעור התצפויות שנמצאות ב"זנבות" של התפלגות – ככלומר, חריגות מהצפי של מודל הדלק לפי סטיית התקן שניתנה (1.34 ס"מ).
- ❖ בטבלה מוצגות: Observed - שיעור תצפויות בפועל שנמצאות מתחת או מעל אחוזנים מסוימים של התפלגות הנורמלית; Expected - שיעור צפוי תיאורטי (לפי התפלגות נורמלית תקנית; רוח סマー (95%) סכיב הערך הצפוי, תחת ההנחה שהמודל תקין).
- ❖ מתחת לאחיזון 3 אנו מצפים לראות % 3 מהעבוררים. בפועל קיבלנו % 8.14 - יותר מפי 2.5 מהצפי, וגם מחוץ לרווח הסマー. גם ב- % 5 Below ו- % 10% יש עודף משמעותי של תצפויות נמוכות ממה שהמודל מנביא.
- ❖ הנתונים מראים הטיה שמאלית באוכלוסייה – ככלומר, מדידות היקף הבטן בישראל נוטות להיות קטנות יותר ממה שמודל הדלק מנביא.

בדיקות שאריות סטנדרטיות מודל הדלק

Standardized Residuals (z-scores) vs Gestational Age



- ❖ בשלב זה חישבנו את השאריות המתוקנות - Z scores (scores) של כל מדידת AC תוך השוואה לערך הצפוי לפי נוסחת הדלק ולפי סטיית התקן.
- ❖ כל נקודה מייצגת מדידת AC של עבור מסוים, לאחר חישוב Z-score
- ❖ הציר האופקי מייצג את שבוע ההריון, והציר האנכי את השארית המתוקנת.
- ❖ הקווים האופקיים מייצגים אחוזונים מרכזים בהתקלות תקינה: קו אדום הוא הצפוי לפי המודל; קווים צבעוניים נוספים: אחוזונים 3%, 5%, 10%, 97.5%, 95%, 90%.
- ❖ ניכרת אסימטריה בrhoה : יש עודף ערכים נמוכים מאוד ($-2 < Z$) כמעט בכל שבועות ההריון.
- ❖ חריגות גבוהות ($2 + > Z$) קיימות, אך שכיחות הרבה פחות.
- ❖ התקנון מאפשר לזהות חריגות מערכתיות ממודל הדלק ומראה כי קיימת נטייה למדידות נמוכות יותר מהצפוי באוכלוסייה הישראלית – ממצא עקי לאורך שבועות ההריון.

ניתוח זנבות במודל Kurmanavicius

Abdominal circumference (AC) - היקף בטן:
 $\text{mean} = 89.39 + 12.03w - 0.000863w^3$

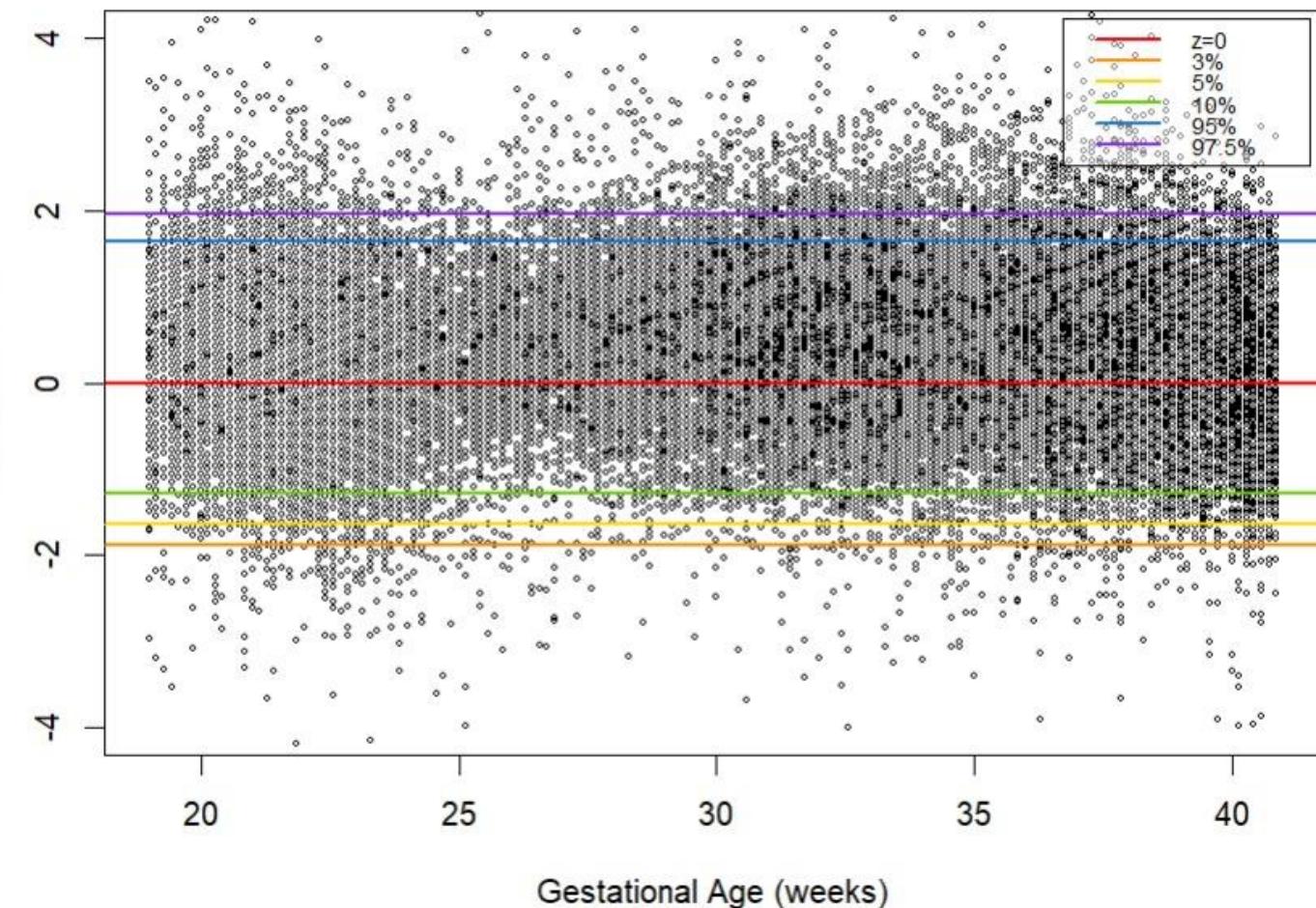
$SD = 1.179 + 0.4753w$

	Tail	Observed	Expected	Lower_CI	Upper_CI
1	Below 3%	0.009199164	0.030	0.02854529	0.03145471
2	Below 5%	0.014689574	0.050	0.04814144	0.05185856
3	Below 10%	0.031815290	0.100	0.09744171	0.10255829
4	Above 95%	0.069884556	0.050	0.04814144	0.05185856
5	Above 97.5%	0.036396691	0.025	0.02366862	0.02633138

- ❖ ביצעו ניתוח של שיעור התצפויות שנפלו בزنבות של התפלגות השאריות (Z-scores) לפי מודל קורמןביצ'וס. המטרה הייתה לבדוק האם הנתונים האמפיריים תואמים את הציפיות מהתפלגות נורמלית תקנית.
- ❖ Below 3%: שיעור התצפויות בפועל הוא %0.92, לעומת ציפייה ל-%3. מובהק סטטיסטית – נמצא מתחת לרוח הסמרק (%-2.85% – 3.14%).
- ❖ Below 5%: שיעור בפועל ~1.47%, ציפייה: %.5. גם כאן, נמצא משמעותית ומוחזק לרוח הסמרק.
- ❖ Below 10%: בפועל ~3.18%, ציפייה: 10%. פער גדול במיוחד, מוחזק לרוח הסמרק (9.74% – 10.26%).
- ❖ Above 95%: שיעור בפועל ~6.99%, גובה מהצפי (%) ומעלה הרף העליון של רוח הסמרק (5.19%) – ככלומר גם עוזף מובהק של ערכים גבוהים.
- ❖ Above 97.5%: שיעור בפועל ~3.64%, בעוד שהצפי הוא %2.5%, ורוח הסמרק נגמר ב-2.63%. גם כאן, מובהקות סטטיסטית כלפי מעלה.
- ❖ בשונה ממודל הדלק שבו נמצאו יותר מדי ערכים נמוכים מהצפי, במודל קורמןביצ'וס מתקבל דפוס הפוך: יש פחות מדי תצפויות נמוכות, יותר מדי תצפויות גבוהות – באופן מובהק סטטיסטית.

בדיקה שאריות סטנדרטיות מודל Kurmanavicius

Standardized Residuals (z-scores) vs Gestational Age



- בשלב זה חישבנו את השאריות המתוקנות (-Z scores) של כל מדידת AC תוך השוואה לערך הצפוי לפי נוסחת Kurmanavicius ולפי סטיית התקן המשוערת לכל שבוע הירionario.
- כל נקודה מייצגת מדידת AC של עובר מסוים לאחר חישוב Z-score הציר האופקי מייצג את שבוע ההריון, והציר האנכי את השארית המתוקנת. הקווים האופקיים מייצגים אחוזונים מרכזיים בהתרפלגות תקינה: קו אדום הוא הצפוי לפיה המודל, והקוויים הצבועוניים מצינים את האחוזונים 97.5%, 95%, 10%, 5%, 3%.
- בניגוד למודל הדלוק, כאן ניכרת אסימטריה הפוכה: יש עודף ערכים חיוביים לאורך כל טווחי הירionario, בעוד שמספר הערכים הנמוכים קטן מהצפוי.
- הדף מעיד כי מדידות AC בנסיבות האמיתיות נוטות להיות גבוהות יותר מהתחזית של קורמןביצ'יס, כלומר המודל מעריך בחסר את היקף הבطن של העוברים באוכלוסייה הישראלית.