

מבוא לסטטיסטיקה ב, קיץ 2025

ביה"ס לכלכלה

המכללה למנהל

סמסטר אביב 2025

לימודיה

רועי עידן



יש פלטי SAS וגם פלטי אקסל

(פלטי STATA לא בחומר)

רוב הנושאים יכולים להופיע ללא או עם פלט

לרוב הנושאים יש גם רווח סמך

הנושאים דומים אחד לשני ועם שלבי עבודה דומים

רשימת הנושאים הסופית למבחן:

1. בדיקת השערות כאשר סטיית התקן באוכלוסייה נתונה לוח Z + רווח סמך. בדף נוסחאות נקרא: בדיקת השערות לתוחלת - שונות ידועה

2. בדיקת השערות כאשר סטיית התקן במדגם S נתונה לוח T + רווח סמך. בדף נוסחאות נקרא: בדיקת השערות לתוחלת - שונות לא ידועה

3. בדיקת השערות על שני מדגמים תלויים (מזווגים), לוח T + רווח סמך. בדף נוסחאות נקרא:

בדיקת השערות ל μ_D - הפרש תוחלות $\mu_1 - \mu_2$ במדגמים תלויים (מזווגים)

4. בדיקת השערות על 2 מדגמים בלתי תלויים:

בדיקת השערות ורווח סמך לשוויון תוחלות (הפרש התוחלות $\mu_1 - \mu_2$) במדגמים בלתי תלויים

בבלתי תלויים, יש 3 מקרים:

א. השונויות באוכלוסיות ידועות (הכוונה נתונות במספרים) + רווח סמך.

ב. השונויות באוכלוסיות לא ידועות (לא נתונות במספרים) אך הן שוות + רווח סמך.

ג. השונויות באוכלוסיות לא ידועות (לא נתונות במספרים) ולא שוות + רווח סמך.

5. מבחן F לבדיקת השערות על שונויות האוכלוסייה בדף נוסחאות נקרא: **בדיקת השערות לשוויון שונויות במדגמים בלתי תלויים**

חובה לתרגל רק מבחנים בשלב זה, לא תרגילי בית!

המבחן כולו פתוח, נקבל דף נוסחאות (לעבוד עימו במידי) + טבלת F, T, Z

יש הרבה חומר, המבחן יצירתי ומתחדש. חשוב להתייחס בשיא הרצינות.

אורך המבחן: שעתיים וחצי
בכל שאלת בדיקת השערות יש לרשום את ההשערות
רמת המובהקות היא תמיד 5% אלא אם נאמר אחרת.

מבנה המבחן:

שאלת היגדים 30 נקודות סך הכל:
6 היגדים מכל נושאי הקורס, 5 נקודות לטענה.
ההסברים חייבים להיות טובים (לא לחפור אלא פשוט לומר את מה שצריך).

70 הנקודות הנותרות:
כל שאר השאלות פתוחות מרובות סעיפים. חלקן עם או בלי פלט.

רועי עידן

מקצוען בהכנת סטודנטים למבחני התואר

מומחה למבחני המכללה למנהל

מפעיל אתר להכנה למבחני המכללה למנהל

באמצעות קורסים מקוונים-

סרטוני הכנה למבחן

ב 16 השנים האחרונות העברתי

את כל הלימודיות הרשמיות בקורסים

סטטיסטיקה א + ב

מתמטיקה א + ב

אקונומטריקה א+ב

בביה"ס לכלכלה, המכללה למנהל

20%
Off

קורס מקוון בסטטיסטיקה ב:

הנחה של 20%

קוד קופון: : אביבים

<https://roy-idan.co.il/hmh8>

קורס מקוון במתמטיקה ב:

הנחה של 20%

קוד קופון: : אביבים

<https://roy-idan.co.il/mb6e>

קורס מקוון בחשבונאות פיננסית ב:

הנחה של 20%

קוד קופון: : אביבים

<https://roy-idan.co.il/pvv6>

שאלות שירדו מהחומר:

תשפ"ב – מועד א –

שאלה 1 סעיף ב
שאלה 4

תשפ"ב – מועד ב –

שאלה 1 סעיף ב
שאלה 3

תשפ"ג – מועד א –

שאלה 1 סעיף ה
שאלה 3

תשפ"ג – מועד ב –

שאלה 1 סעיף ה
שאלה 3

תשפ"ד – מועד ב –

שאלה 1 סעיף א

מבחני תשפ"ה 2025 הם הכי עדכניים וחשובים:

צריך לתרגל את כולם!!!

חומר שירד:

פלט של תוכנת STATA:

Two-sample t test with equal variances

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
no	61	13.11984	1.639066	12.80151	9.841216	16.39846
yes	106	8.281604	.4846764	4.990049	7.32058	9.242628
combined	167	10.04886	.6940447	8.969034	8.67857	11.41915
diff					2.082221	7.594243

diff = mean(**no**) - mean(**yes**)
Ho: diff = 0
Ha: diff < 0
Pr(T < t) = **0.9997**
Ha: diff != 0
Pr(|T| > |t|) = **0.0007**
Ha: diff > 0
Pr(T > t) = **0.0003**
t = **3.4662**
degrees of freedom = **165**

Two-sample t test with unequal variances

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
no	61	13.11984	1.639066	12.80151	9.841216	16.39846
yes	106	8.281604	.4846764	4.990049	7.32058	9.242628
combined	167	10.04886	.6940447	8.969034	8.67857	11.41915
diff					1.429837	8.246628

diff = mean(**no**) - mean(**yes**)
Ho: diff = 0
Ha: diff < 0
Pr(T < t) = **0.9970**
Ha: diff != 0
Pr(|T| > |t|) = **0.0060**
Ha: diff > 0
Pr(T > t) = **0.0030**
t = **2.8307**
Satterthwaite's degrees of freedom = **70.6429**

Variance ratio test

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
no	61	13.11984	1.639066	12.80151	9.841216	16.39846
yes	106	8.281604	.4846764	4.990049	7.32058	9.242628
combined	167	10.04886	.6940447	8.969034	8.67857	11.41915

ratio = sd(**no**) / sd(**yes**)
Ho: ratio = 1
Ha: ratio < 1
Pr(F < f) = **1.0000**
Ha: ratio != 1
2*Pr(F > f) = **0.0000**
Ha: ratio > 1
Pr(F > f) = **0.0000**
f = **6.5813**
degrees of freedom = **60, 105**

חומר שירד- כל מה שקשור להתפלגות נורמלית (תרגיל 1 מהמודל):

תרגיל 1 - התפלגות נורמלית

1. המשתנה X - הגובה של גברים בגיל $18+$

נתון: $X \sim N(175, 16)$

א. מהו אחוז הגברים שגובהם מעל 180 ס"מ

ב. מהו גובהו של גבר ש - 80% מהגברים נמוכים ממנו

2. מכונה אוטומטית, למילוי כוסות במשקה קל, מכוונת למלא במוצק 200 גרם לכוס.

אם כמות המשקה בכוס מתפלגת נורמלית עם סטיית תקן של 6 גרם

א. מהו אחוז הכוסות המכילות לכל היותר 196 גרם של המשקה?

ב. מהו אחוז הכוסות המכילות בין 192 גרם ל - 208 גרם של המשקה?

ג. מהם התחום הסימטרי סביב התוחלת בו מרוכזים 90% מהכוסות?

3. $X =$ ציוני המבחן הפסיכומטרי לכניסה למוסדות להשכלה גבוהה $X \sim N(500, 120^2)$

אם נבחר סטודנט באקראי

א. מה ההסתברות שציונו יהיה לפחות 510

ב. מה ההסתברות שציונו יהיה בין 495 לבין 508

ג. מהם התחום הסימטרי סביב התוחלת בו מרוכזים 95% מציוני הסטודנטים

אם נבחר מדגם של 225 סטודנטים

א. מה ההסתברות שממוצע המדגם הנבחר יהיה לפחות 510

ב. מה ההסתברות שממוצע המדגם הנבחר יהיה בין 495 לבין 508

ג. מהו התחום הסימטרי סביב התוחלת בו מרוכזים 95% מכל המדגמים האפשריים בגודל $n=225$

4. קוטר של כבל חשמלי מתפלג נורמלית עם ממוצע של 10 מ"מ וסטיית תקן של 0.2 מ"מ.

כבל נחשב פגום אם קוטרו מעל 10.3 מ"מ.

מהו אחוז הכבלים הפגומים הצפוי בתהליך הייצור?

טבלאות שנקבל במבחן:

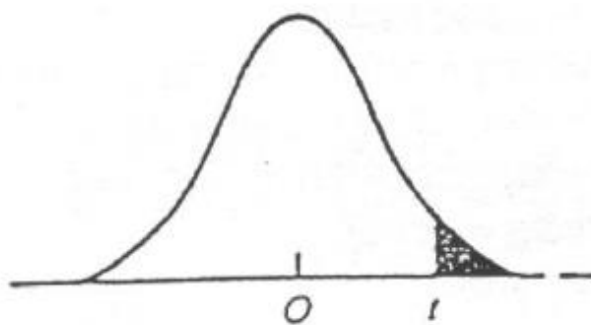
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

z	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.891	4.417
$\Phi(z)$	0.90	0.95	0.975	0.99	0.995	0.999	0.9995	0.99995	0.999995

רועי עידן- מומחה להכנת סטודנטים למבחני המכללה למנהל: כלכלה, סטטיסטיקה, מתמטיקה
 סרטוני הכנה למבחני המכללה למנהל (קורסים מקוונים): www.roy-idan.co.il 052-546-6016

Table B-2 Student's t distribution

The first column lists the number of degrees of freedom (ν). The headings of the other columns give probabilities (P) for t to exceed the entry value. Use symmetry for negative t values.



$\nu \backslash P$	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

Reprinted from P. B. Hoel, *Introduction to Mathematical Statistics*, 4th ed., New York, Wiley, 1971, by permission of the publishers.

טבלה 4. טבלת התפלגות F ערכי החלוקה F_p של התפלגות $F(m, n)$

m — דרגות חופש המונה; n — דרגות חופש המכונה

p	n	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	120	∞
.95	1	161	200	216	225	230	234	237	239	241	242	244	246	248	250	252	253	254
.975		648	800	864	900	922	937	948	957	963	969	977	985	993	1001	1010	1014	1018
.99		4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6106	6157	6209	6261	6313	6339	6366
.995		16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426	24630	24836	25044	25253	25359	25464
.95	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.46	19.48	19.49	19.50
.975		38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.48	39.49	39.50
.99		98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.42	99.43	99.45	99.47	99.48	99.49	99.50
.995		198.50	199.01	199.16	199.24	199.30	199.33	199.36	199.38	199.39	199.39	199.42	199.43	199.45	199.48	199.48	199.49	199.51
.95	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.62	8.57	8.55	8.53
.975		17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.08	13.99	13.95	13.90
.99		34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.05	26.87	26.69	26.50	26.32	26.22	26.13
.995		55.55	49.80	47.47	46.20	45.39	44.84	44.43	44.13	43.88	43.68	43.39	43.08	42.78	42.47	42.15	41.99	41.83
.95	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.69	5.66	5.63
.975		12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.46	8.36	8.31	8.26
.99		21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.84	13.65	13.56	13.46
.995		31.33	26.28	24.26	23.15	22.46	21.98	21.62	21.35	21.14	20.97	20.70	20.44	20.17	19.89	19.61	19.47	19.32
.95	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.43	4.40	4.37
.975		10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.23	6.12	6.07	6.02
.99		16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.38	9.20	9.11	9.02
.995		22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.38	13.15	12.90	12.66	12.40	12.27	12.14
.95	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.74	3.70	3.67
.975		8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.07	4.96	4.90	4.85
.99		13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.23	7.06	6.97	6.88
.995		18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.03	9.81	9.59	9.36	9.12	9.00	8.88
.95	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.38	3.30	3.27	3.23
.975		8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.36	4.25	4.20	4.14
.99		12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	5.99	5.82	5.74	5.65
.995		16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.18	7.97	7.75	7.53	7.31	7.19	7.08
.95	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	3.01	2.97	2.93
.975		7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.89	3.78	3.73	3.67
.99		11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.20	5.03	4.95	4.86
.995		14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01	6.81	6.61	6.40	6.18	6.06	5.95

m

p	n	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	120	∞
.95	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.79	2.75	2.71
.975		7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.56	3.45	3.39	3.33
.99		10.56	8.02	6.89	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.65	4.48	4.40	4.31
.995		13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23	6.03	5.83	5.62	5.41	5.30	5.19
.95	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.70	2.62	2.58	2.54
.975		6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.31	3.20	3.14	3.08
.99		10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.25	4.08	4.00	3.91
.995		12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66	5.47	5.27	5.07	4.86	4.75	4.64
.95	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.47	2.38	2.34	2.30
.975		6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	2.96	2.85	2.79	2.72
.99		9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.70	3.54	3.45	3.36
.995		11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91	4.72	4.53	4.33	4.12	4.01	3.90
.95	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.16	2.11	2.07
.975		6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.64	2.52	2.46	2.40
.99		8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.21	3.05	2.96	2.87
.995		10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25	4.07	3.88	3.69	3.48	3.37	3.26
.95	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.95	1.90	1.84
.975		5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.35	2.22	2.16	2.09
.99		8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.78	2.61	2.52	2.42
.995		9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.68	3.50	3.32	3.12	2.92	2.81	2.69
.95	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.74	1.68	1.62
.975		5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.07	1.94	1.87	1.79
.99		7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.39	2.21	2.11	2.01
.995		9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.18	3.01	2.82	2.63	2.42	2.30	2.18
.95	60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.65	1.53	1.47	1.39
.975		5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.82	1.67	1.58	1.48
.99		7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.03	1.84	1.73	1.60
.995		8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.74	2.57	2.39	2.19	1.96	1.83	1.69
.95	120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.55	1.43	1.35	1.25
.975		5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.69	1.53	1.43	1.31
.99		6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.86	1.66	1.53	1.38
.995		8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54	2.37	2.19	1.98	1.75	1.61	1.43
.95	∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.32	1.22	1.00
.975		5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.57	1.39	1.27	1.00
.99		6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.70	1.47	1.32	1.00
.995		7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36	2.19	2.00	1.79	1.53	1.36	

דף הנוסחאות שנקבל במבחן:

דף נוסחאות מבוא לסטטיסטיקה ב'

~~התפלגות נורמלית (התפלגות גאוס)~~

~~$P(Z \leq a) = \Phi(a)$~~

~~$P(Z \geq a) = 1 - P(Z \leq a) = 1 - \Phi(a)$~~

~~$P(Z \leq -a) = \Phi(-a) = 1 - \Phi(a)$~~

~~$P(a \leq Z \leq b) = P(Z \leq b) - P(Z \leq a) = \Phi(b) - \Phi(a)$~~

בדיקת השערות לתוחלת μ - שונות ידועה

$H_0: \mu = \mu_0 \quad (\mu \leq \mu_0)$ $H_1: \mu > \mu_0$ נדחה H_0 כאשר: $Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > Z_{1-\alpha}$ הערך הקריטי מלוח Z הסטטיסטי המחושב במדגם $Pvalue = (\alpha') = 1 - \Phi(Z_{\bar{X}})$	$H_0: \mu = \mu_0 \quad (\mu \geq \mu_0)$ $H_1: \mu < \mu_0$ נדחה H_0 כאשר: $Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} < -Z_{1-\alpha}$ הערך הקריטי מלוח Z הסטטיסטי המחושב במדגם $Pvalue = (\alpha') = 1 - \Phi(Z_{\bar{X}})$	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$ נדחה H_0 כאשר: $ Z = \left \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \right > Z_{1-\frac{\alpha}{2}}$ הערך הקריטי מלוח Z הסטטיסטי המחושב במדגם $Pvalue = (\alpha') = 2 \cdot (1 - \Phi(Z_{\bar{X}}))$
--	---	---

רווח סמך ל μ ברמת סמך $1 - \alpha$ (כאשר σ ידועה)

$$\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

אורך רווח הסמך $L = L_2 - L_1 = 2 \cdot Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

הסטייה המירבית (המרחק המירבי) של $\bar{X}$ מ- μ $\frac{L}{2} = \bar{X} - \mu = Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

מציאת גודל המדגם n

$$n \geq \frac{\left(Z_{1-\frac{\alpha}{2}} \cdot \sigma \right)^2}{\left(\frac{L}{2} \right)^2}$$

בדיקת השערות לתוחלת μ - שונות לא ידועה (כאשר σ לא ידועה)

$H_0: \mu = \mu_0 \quad (\mu \leq \mu_0)$ $H_1: \mu > \mu_0$ נדחה H_0 כאשר: $t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} > t_{(n-1, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם	$H_0: \mu = \mu_0 \quad (\mu \geq \mu_0)$ $H_1: \mu < \mu_0$ נדחה H_0 כאשר: $t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} < -t_{(n-1, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$ נדחה H_0 כאשר: $ t = \left \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} \right > t_{(n-1, 1-\frac{\alpha}{2})}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם
$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$ כאשר		
רווח סמך ל μ ברמת סמך $1 - \alpha$ $\bar{X} \pm t_{(n-1, 1-\frac{\alpha}{2})} \cdot \frac{S}{\sqrt{n}}$		

בדיקת השערות ל μ_D - הפרש תוחלות $\mu_1 - \mu_2$ במדגמים תלויים (מזווגים)

$H_0: \mu_D = a \quad (\mu_D \leq a)$ $H_1: \mu_D > a$ נדחה H_0 כאשר: $t = \frac{\bar{d} - a}{\frac{S_d}{\sqrt{n}}} > t_{(n-1, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם	$H_0: \mu_D = a \quad (\mu_D \geq a)$ $H_1: \mu_D < a$ נדחה H_0 כאשר: $t = \frac{\bar{d} - a}{\frac{S_d}{\sqrt{n}}} < -t_{(n-1, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם	$H_0: \mu_D = a$ $H_1: \mu_D \neq a$ נדחה H_0 כאשר: $ t = \left \frac{\bar{d} - a}{\frac{S_d}{\sqrt{n}}} \right > t_{(n-1, 1-\frac{\alpha}{2})}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם
$S = S_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}$		
רווח הסמך ל μ_D ברמת סמך $1 - \alpha$ $\bar{d} \pm t_{(n-1, 1-\frac{\alpha}{2})} \cdot \frac{S}{\sqrt{n}}$		

בדיקת השערות ורווח סמך לשוויון תוחלות (הפרש התוחלות $\mu_1 - \mu_2$) במדגמים בלתי תלויים

שונויות ידועות $\sigma_1^2, \sigma_2^2 \Leftarrow$ ידועות

$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 > a$ נדחה H_0 כאשר: $Z = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > Z_{1-\alpha}$ הערך הקריטי מלוח Z הסטטיסטי המחושב במדגם $Pvalue = (\alpha') = 1 - \Phi(Z_{\bar{X}_1 - \bar{X}_2})$	$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 < a$ נדחה H_0 כאשר: $Z = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < -Z_{1-\alpha}$ הערך הקריטי מלוח Z הסטטיסטי המחושב במדגם $Pvalue = (\alpha') = 1 - \Phi(Z_{\bar{X}_1 - \bar{X}_2})$	$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 \neq a$ נדחה H_0 כאשר: $ Z = \frac{ \bar{X}_1 - \bar{X}_2 - a }{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > Z_{1-\frac{\alpha}{2}}$ הערך הקריטי מלוח Z הסטטיסטי המחושב במדגם $Pvalue = (\alpha') = 2 \cdot (1 - \Phi(Z_{\bar{X}_1 - \bar{X}_2}))$
רווח סמך ל $\mu_1 - \mu_2$ ברמת סמך $1 - \alpha$ $(\bar{X}_1 - \bar{X}_2) \pm Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$		

שונויות לא ידועות, אך שוות $\sigma_1^2 = \sigma_2^2 = \sigma^2 \Leftarrow$ שוות

$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 > a$ נדחה H_0 כאשר: $t = \frac{\bar{X}_1 - \bar{X}_2 - a}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_{(n_1+n_2-2, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי המחושב במדגם	$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 < a$ נדחה H_0 כאשר: $t = \frac{\bar{X}_1 - \bar{X}_2 - a}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < -t_{(n_1+n_2-2, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי המחושב במדגם	$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 \neq a$ נדחה H_0 כאשר: $ t = \frac{ \bar{X}_1 - \bar{X}_2 - a }{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_{(n_1+n_2-2, 1-\frac{\alpha}{2})}$ הערך הקריטי מלוח t הסטטיסטי המחושב במדגם
$S^2 = \frac{\sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2 + \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2}{n_1 + n_2 - 2} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$		
רווח סמך ל $\mu_1 - \mu_2$ ברמת סמך $1 - \alpha$ $(\bar{X}_1 - \bar{X}_2) \pm t_{(n_1+n_2-2, 1-\frac{\alpha}{2})} \cdot S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$		

שונויות לא ידועות ולא שוות $\sigma_1^2 \neq \sigma_2^2 \Leftrightarrow$

$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 > a$ נדחה H_0 כאשר: $t = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} > t_{(df, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם	$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 < a$ נדחה H_0 כאשר: $t = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} < -t_{(df, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם	$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 \neq a$ נדחה H_0 כאשר: $ t = \left \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \right > t_{(df, 1-\frac{\alpha}{2})}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם
$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} \right)^2}{\frac{\left(\frac{S_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2} \right)^2}{n_2 - 1}}$		
רווח סמך ל $\mu_1 - \mu_2$ ברמת סמך $1 - \alpha$ $(\bar{X}_1 - \bar{X}_2) \pm t_{(df, 1-\frac{\alpha}{2})} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$		

בדיקת השערות לשוויון שונויות ($\sigma_1^2 = \sigma_2^2$) במדגמים בלתי תלויים

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 > \sigma_2^2$$

נדחה H_0 כאשר:

$$F = \frac{\frac{\sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2}{n_1 - 1}}{\frac{\sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2}{n_2 - 1}} > F(n_1 - 1, n_2 - 1, 1 - \alpha)$$

הערך הקריטי
מלוח F

הסטטיסטי F המחושב
במדגם

המורכבות הגדולה בקורס: "שני מדגמים בלתי תלויים"

בנושא "שני מדגמים בלתי תלויים" יש 3 מקרים:

כל אחד מהם עשוי להגיע עם או ללא פלט.

מקרה א:

השונות באוכלוסיות ידועות (נתונות במספרים).

מבחן Z

מקרה ב:

השונות באוכלוסיות לא ידועות (לא נתונות במספרים) אך הן שוות.

מבחן T

מקרה ג:

השונות באוכלוסיות לא ידועות (לא נתונות במספרים) אך הן לא שוות.

מבחן T

2 מדגמים בלתי תלויים,

שוניות האוכלוסיות נתונות (ידועות): מקרה א בדף נוסחאות:

בדיקת השערות ורווח סמך לשוויון תוחלות (הפרש התוחלות $\mu_1 - \mu_2$) במדגמים בלתי תלויים

$$\text{שוניות ידועות} \Leftarrow \sigma_1^2, \sigma_2^2$$

על מנת לבדוק את הטענה שתוחלת שטח הדירה ירד עם השנים, נלקחו נתונים

על 200 דירות בשנת 2019 והתקבל שטח דירה ממוצע של 112.5 מ"ר

ונתונים על 150 דירות בשנת 2009 והתקבל שטח דירה ממוצע של 120 מ"ר

ידוע כי סטיית התקן של שטח הדירה בשנת 2019 הייתה 30 מ"ר ובשנת 2009 סטיית התקן הייתה 20 מ"ר

נרכז את הנתונים:

2019	2009
$N_1=200$	$N_2=150$
$112.5 = \bar{x}_1$	$120 = \bar{x}_2$
$30 = \sigma_1$	$20 = \sigma_2$

(10) א. בידקו את הטענה ברמת מובהקות של 1%

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 < 0$$

נדחה H_0 כאשר:

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < \underbrace{-Z_{1-\alpha}}$$

הסטטיסטי Z
המחושב במדגם

הערך הקריטי
 Z מלוח

אכן דוחים H_0

מקבלים H_1 תוחלת שטח הדירה ירד עם השנים וזאת בר"מ 1%

חוק:

בכל נושאי הקורס, לחישוב PV תמיד נפעיל את הנוסחה המרכזית של אותו נושא מדף הנוסחאות, אשר תמיד תוצאתו תמיד נקראת סטטיסטי

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

הסטטיסטי Z
המחושב במדגם

תמיד חובה להתייחס ל Z כפלוס (כי בטבלת Z יש רק Z ים חיוביים)

$$Pvalue = (\alpha') = 1 - \Phi(|Z_{\bar{X}_1 - \bar{X}_2}|)$$

החוק הכי חשוב בקורס (לשנן):

אם רמת מובהקות $p_v <$

אז דוחים H_0 , מקבלים H_1

אם רמת מובהקות $p_v >$

אז מקבלים H_0 , דוחים H_1

(5) ג. חשבו רווח סמך ברמת סמך של 99%

רמת מובהקות אלפא = 1% 0.01

רווח סמך ל $\mu_1 - \mu_2$ ברמת סמך $1 - \alpha$

$$(\bar{X}_1 - \bar{X}_2) \pm Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$Z_{1-\frac{\alpha}{2}}$$

$$(\bar{X}_1 - \bar{X}_2) - 2.576 \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad , \quad (\bar{X}_1 - \bar{X}_2) + 2.576 \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

(3) ד. מהי משמעות רווח הסמך שחישבתם בסעיף הקודם

ברמת סמך של 99%

תוחלת שטח דירה ב 2019 נמוכה לעומת שטח דירה ב 2009

בין 0.64 לבין 14.39 מ"ר

(5) ה. התברר כי המדגם בשנת 2009 היה של 300 תצפיות והמדגם של 2019 של 450 תצפיות.
שאר הנתונים ללא שינוי. האם ניתן ללא חישוב נוסף לדעת מה תהיה החלטת המבחן בסעיף א'.
אם כן, הציגו את ההחלטה ואם לא הסבירו מדוע

תזכורת לחוק:

אם רמת מובהקות $p_v <$

אז דוחים H_0 , מקבלים H_1

בסעיף א דחינו H_0 לכן נסיק שה PV קטנה מרמת מובהקות 1%

חוק:

כאשר N גדל ה PV תמיד קטן

הרי במצב המוצא ה PV היה קטן מר"מ של 1%

אחרי ש N גדל, ה PV קטן לכן

עדיין ה PV החדש והמוקטן עדיין נמוך מר"מ 1% לכן שוב דוחים את H_0

כעת שאלת היגד (טענה) באותו נושא:

שני מדגמים בלתי תלויים, שונויות אוכלוסיה נתונות, מקרה א:

ב. על מנת לבדוק את ההשערה שיש הבדל בתוחלת שנות הלימוד בין שחורים ולבנים הוצג הפלט הבא

z-Test: Two Sample for Means		
	שנות לימוד לבנים	שנות לימוד שחורים
Mean	13.77655476	12.96793003
Known Variance	55	55
Observations	1817	343
Hypothesized Mean Difference	0	
z	1.852095273	
P(Z<=z) one-tail		
z Critical one-tail		
P(Z<=z) two-tail		
z Critical two-tail		

נכון/לא נכון/אי אפשר לדעת

Pvalue לבדיקת הטענה הינו 0.0322

ניקח את Z סטטיסטי מהפלט: 1.85

ובטבלת Z, מבחוץ לבפנים נביא שטח של 0.9678

$$Pvalue = (\alpha') = 2 \cdot (1 - \Phi(|Z_{\bar{X}_1 - \bar{X}_2}|))$$

ההיגד שגוי

בדיקת השערות ורווח סמך לשוויון תוחלות (הפרש התוחלות $\mu_1 - \mu_2$) במדגמים בלתי תלויים
שונות לא ידועות, אך שוות $\sigma_1^2 = \sigma_2^2 = \sigma^2 \Leftarrow$ מקרה ב בדף נוסחאות:

מועד ב 2024:

שאלה 3 (25 נקודות)

על מנת לבדוק את הטענה שנשים משכילות יותר מגברים (תוחלת שנות הלימוד של נשים גבוהה מזו של גברים), הוצג הפלט הבא :

t-Test: Two-Sample Assuming Equal Variances

	women	men
Mean	14	11
Variance	37.81	31.49
Observations	93	176
Pooled Variance	33.66	
Hypothesized Mean Difference	0	
df	267	
t Stat		
P(T<=t) one-tail	3.47652E-05	
t Critical one-tail		
P(T<=t) two-tail	6.95303E-05	
t Critical two-tail		

t-Test: Two-Sample Assuming Unequal Variances

	women	men
Mean	14	11
Variance	37.81	31.49
Observations	93	176
Hypothesized Mean Difference	0	
df	173	
t Stat		
P(T<=t) one-tail	6.1607E-05	
t Critical one-tail		
P(T<=t) two-tail	0.000123214	
t Critical two-tail		

F-Test Two-Sample for Variances

	women	men
Mean	14	11
Variance	37.81	31.49
Observations	93	176
df	92	175
F		
P(F<=f) one-tail	0.151358651	
F Critical one-tail		

(8) א. בדקו את ההשערה ללא חישובים נוספים

(הגדירו את ההשערות המתאימות לכל מבחן שאתם עורכים)

קודם נבצע מבחן F לשוויון שונויות:

מבחן F לשוויון שונויות

מקרה ב בדף נוסחאות $H_0: \sigma_1^2 = \sigma_2^2$

מקרה ג בדף נוסחאות $H_1: \sigma_1^2 > \sigma_2^2$

מסקנה: נקבל H_0 השונויות שוות.

אנו במקרה ב של דף הנוסחאות.

רק אחרי שגילינו באיזה נושא אנו נמצאים, נוכל לבצע את מבחן T

לבדיקת הטענה מהשאלה: "נשים משכילות יותר מגברים":

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 > 0$$

מסקנה: דוחים השערת האפס

השערת החוקר נכונה: נשים משכילות יותר מגברים וזאת בר"מ 5%

(8) ב. בדקו את הטענה שתוחלת שנות הלימוד של נשים גבוהה מזו של הגברים ביותר מ 2 שנות לימוד

חוק לכל הקורס:

מותר לעבוד עם PV מהפלט רק אם בהשערות יש את המספר 0

כתיבת השערות:

$$H_0: \mu_1 - \mu_2 = 2$$

$$H_1: \mu_1 - \mu_2 > 2$$

$$t = \frac{\bar{X}_1 - \bar{X}_2 - a}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{14 - 11 - 2}{\sqrt{33.66} \sqrt{\frac{1}{93} + \frac{1}{176}}} = \frac{3 - 2}{0.743} = 1.34$$

טי סטטיסטי (תוצאת הנוסחה) הוא 1.34

כעת נביא מטבלת T את ה T הקריטי:

$$t_{(267, 0.95)} =$$

מתוך דף הנוסחאות:

נדחה H_0 כאשר:

$$t = \frac{\bar{X}_1 - \bar{X}_2 - a}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_{(n_1+n_2-2, 1-\alpha)}$$

הסטטיסטי t
המחושב
במדגם

הערך הקריטי
מלוח t

מסקנה:

הטי הסטטיסטי קטן מהקריטי לכן נקבל H_0 (לא דוחים H_0)

נדחה H_1

תוחלת שנות לימוד נשים לא גבוהה משל הגברים ביותר מ 2 שנות לימוד

S

בסעיף ג "זורקים" אותנו לנושא אחר (הפעם מדגם יחיד, מבחן T):

בדיקת השערות לתוחלת μ - שונות לא ידועה (כאשר σ לא ידועה)

(5) ג. בדקו את הטענה כי הגברים בישראל משכילים יותר מגברים בארה"ב. ידוע שתוחלת שנות לימוד של גברים בארה"ב היא 10 שנות לימוד.

$$H_0: \mu_2 = 10$$

$$H_1: \mu_2 > 10$$

$$t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{11 - 10}{\frac{\sqrt{31.49}}{\sqrt{176}}} = \frac{1}{0.422} = 2.369$$

טי סטטיסטי (תוצאת הנוסחה) הוא 2.369

הבאת T קריטי מטבלת T:

$$t_{(175, 0.95)} =$$

מתוך דף הנוסחאות:

$$H_0 : \mu = \mu_0 \quad (\mu \leq \mu_0)$$

$$H_1 : \mu > \mu_0$$

נדחה H_0 כאשר:

$$t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} > t_{(n-1, 1-\alpha)}$$

הסטטיסטי
המחושב
במדגם

תערך הקריטי
מלוח t

הסטטיסטי מהנוסחה גדול מהקריטי שבטבלה:

לכן, נדחה H_0

נקבל H_1

השכלת גברים בישראל אכן גבוהה מארה"ב, בר"מ 5%

(4) ד. השלימו את המספר המסומן בריבוע

$$F = \frac{S_1^2}{S_2^2} = \frac{37.81}{31.49} = 1.2$$

בדיקת השערות ורווח סמך לשוויון תוחלות (הפרש התוחלות $\mu_1 - \mu_2$) במדגמים בלתי תלויים
שונויות לא ידועות ולא שוות $\sigma_1^2 \neq \sigma_2^2 \Leftrightarrow$ מקרה ג בדף נוסחאות:

ממבחן תשפ"ד 2024 שאלה 2

שאלה 2 (35 נקודות)

על מנת לבדוק את הטענה שגברים עובדים יותר שעות מנשים, נדגמו 500 גברים ו 500 נשים.
התקבל שממוצע שעות העבודה של הגברים הוא 220 שעות בחודש עם שונות של 400 שעות
וממוצע שעות העבודה החודשיות של נשים הוא 190 עם שונות של 225 שעות.

גברים	נשים
$N1=500$	$N2=500$
$220 = \overline{x_1}$	$190 = \overline{x_2}$
$400 = S_1^2$ שונות מדגם	$225 = S_2^2$ שונות מדגם

(13) א. בדקו את הטענה ברמת מובהקות של 0.5%
אנחנו בנושא של שני מדגמים בלתי תלויים.
כדי לדעת אם בדף הנוסחאות אנו
במקרה ב (שונויות שוות) או במקרה ג (שונויות לא שוות)
יהיה עלינו בשלב הראשון לערוך מבחן F:

מבחן F לשוויון שונויות	
$H_0: \sigma_1^2 = \sigma_2^2$	מקרה ב בדף נוסחאות
$H_1: \sigma_1^2 > \sigma_2^2$	מקרה ג בדף נוסחאות

$$F = \frac{S_1^2}{S_2^2}$$

F סטטיסטי מהנוסחה הינו 1.77

הבאת F קריטי מלוח F:

$$F(n_1 - 1, n_2 - 1, 1 - \alpha)$$

הערך הקריטי
F מלוח

$$F(499, 499, 0.995) =$$

נדחה H_0 כאשר:

$$F = \frac{S_1^2}{S_2^2} = \frac{\frac{\sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2}{n_1 - 1}}{\frac{\sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2}{n_2 - 1}} > F(n_1 - 1, n_2 - 1, 1 - \alpha)$$

הסטטיסטי F המחושב במדגם

הערך הקריטי F מלוח

אצלינו,

F סטטיסטי מהנוסחה $<$ קריטי מטבלת F לכן:

נדחה H_0

השונויות לא שוות לכן **אנו במקרה ג של דף הנוסחאות.**

רק אחרי שגילינו באיזה נושא אנו (מקרה ג), נוכל לבצע את מבחן T שיבדוק את הטענה מהשאלה: "גברים עובדים יותר שעות מנשים".

שונויות לא ידועות ולא שוות $\sigma_1^2 \neq \sigma_2^2 \Leftrightarrow$

$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 > a$ נדחה H_0 כאשר: $t = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} > t_{(df, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם	$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 < a$ נדחה H_0 כאשר: $t = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} < -t_{(df, 1-\alpha)}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם	$H_0: \mu_1 - \mu_2 = a$ $H_1: \mu_1 - \mu_2 \neq a$ נדחה H_0 כאשר: $ t = \frac{ \bar{X}_1 - \bar{X}_2 - a }{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} > t_{(df, 1-\frac{\alpha}{2})}$ הערך הקריטי מלוח t הסטטיסטי t המחושב במדגם
---	--	--

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$

טי סטטיסטי מהנוסחה:

$$|t| = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{220 - 190 - 0}{\sqrt{\frac{400}{500} + \frac{225}{500}}} = \frac{30}{1.118} = 26.832$$

הבאת T קריטי מטבלת T:

ראשית, נחשב עם נוסחה מפלצתית את דרגות חופש:

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{1.5625}{0.001688} = 925.55$$

נעגל את דרגות החופש ל 925 (תמיד נעגל מטה)

כעת נוכל להביא T קריטי מטבלת T:

$$t_{(925, 0.995)} =$$

טי סטטיסטי 26.83 גדול מ T קריטי 2.576
אז דוחים H0 ומקבלים H1 כלומר טענת החוקר נכונה:
גברים עובדים יותר מנשים, כל זאת בר"מ של 0.5%

(4) ב. חשבו את ה Pvalue לבדיקת הטענה

נזכיר חוק: תמיד PV נחשב עם תוצאת נוסחת הסטטיסטי

מנוסחת הסטטיסטי: T סטטיסטי 26.83

דרגות החופש 925

לכן $P_v < 0.005$

(5) ג. חשבו רווח סמך ברמת סמך של 98%

$$\bar{X}_1 - \bar{X}_2 \pm t_{(925, 0.99)} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$30 \pm 2.326 \cdot 1.118$$

$$30 \pm 2.6$$

$$(27.4, 32.6)$$

(4) ד. מהי משמעות רווח הסמך שחישבתם בסעיף הקודם

ברמת סמך 98%

תוחלת שעות העבודה של גברים גבוהה מזו של נשים בין 27.4 ל 32.6 שעות בחודש

(5) ה. בדקו את הטענה כי נשים עובדות לפחות 20 שעות פחות בחודש לעומת הגברים.

כדאי להפוך, כלומר: גברים עובדים יותר מהנשים לפחות ב 20 שעות

$$H_0: \mu_1 - \mu_2 = 20$$

$$H_1: \mu_1 - \mu_2 > 20$$

T סטטיסטי מהנוסחה:

$$t = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{220 - 190 - 20}{\sqrt{\frac{400}{500} + \frac{225}{500}}} = \frac{10}{1.118} = 8.94$$

T קריטי מטבלת T:

$$t_{(925, 0.95)} = 1.645$$

טי סטטיסטי גדול מטי קריטי לכן דוחים H_0 ומקבלים H_1

גברים עובדים יותר מהנשים לפחות ב 20 שעות ברמת מובהקות 5%

(4) ו. בדקו את הטענה כי הנשים בישראל עובדות פחות שעות מהנשים באירופה שעובדות בממוצע 220 שעות בחודש, ברמת מובהקות של 10%.

$$H_0: \mu_2 = 220$$

$$H_1: \mu_2 < 220$$

T סטטיסטי מהנוסחה:

$$t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} = \frac{190 - 220}{\frac{15}{\sqrt{500}}} = -44.7$$

T קריטי מטבלת T:

$$-t_{(499, 0.9)} = -1.282$$

סטטיסטי קטן מקריטי לכן דוחים H_0 ומקבלים H_1

נשים בישראל עובדות פחות מנשות אירופה, כל זאת בר"מ 10%

מבחן מה 23.7.23 תשפ"ג(הפתרון המלא להורדה מה MOODLE):

שאלה 2 (34 נקודות)

חוקר רוצה לבדוק את האם יש הבדל בשכר בין גברים נשואים לגברים רווקים.
לשם כך אסף נתונים על שכרם של 60 גברים נשואים ו- 80 גברים רווקים .
התקבל שהגברים הנשואים הרוויחו בממוצע 18,000 ₪ בחודש עם סטיית תקן של 2000 ₪
והגברים רווקים הרוויחו בממוצע 17,000 ₪ עם סטיית תקן של 3000 ₪

(13) א. בדקו את הטענה ברמת מובהקות של 1%

(4) ב. חשבו את ה Pvalue לבדיקת הטענה

(5) ג. חשבו רווח סמך ברמת סמך של 90%

(3) ד. מהי משמעות רווח הסמך שחישבתם בסעיף הקודם

(5) ה. בדקו את הטענה כי תוחלת השכר של גברים רווקים נמוך מתוחלת השכר של נשים רווקות
העומדת על 20,000 ₪.

(4) ו. החוקרת טענה כי שגם אם סטיות התקן שחושבו במדגם היו זהות לסטיות התקן באוכלוסייה, תוצאות
המבחן בסעיף א' לא היו משתנות. חוו **ללא חישוב נוסף** את דעתכם על הטענה

השאלה הבאה היא ממועד ב 2022: הפעם מדובר בפלט SAS:

שאלה 4 (32 נקודות)

על מנת לבדוק האם יש הבדל בתוחלת השכר בין נשים לגברים הוצגו התוצאות הבאות ע"י תוכנת SAS מתוך נתוני סקר הכנסות 2015 מהלשכה המרכזית לסטטיסטיקה

Women=0 מייצג גברים

Women=1 מייצג נשים

The SAS System

The TTEST Procedure

Variable: wage (wage)

women	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		9298	8862.6	8628.0	89.4780	51.0000	173468
1		8197	5498.9	5220.3	57.6594	50.0000	157666

women	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
0		8862.6	8687.2 9038.0	8628.0	8505.8 8753.9
1		5498.9	5385.9 5612.0	5220.3	5141.6 5301.5

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	17493	30.69	<.0001
Satterthwaite	Unequal	15575	31.60	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	9297	8196	2.73	<.0001

(7) א. בדקו ללא חישובים נוספים את הטענה ברמת מובהקות של 1%.
(הגדירו את ההשערות המתאימות לכל מבחן שאתם עורכים)

ראשית, נערוך מבחן F לגבי השונויות כדי לדעת האם הן שוות (מקרה ב)

או שאולי הן לא שוות (מקרה ג). אנו רוצים לברר באיזה נושא אנחנו!

סעיף א:

מבחן F לשוויון שונויות

מקרה ב בדף נוסחאות $H_0: \sigma_1^2 = \sigma_2^2$

מקרה ג בדף נוסחאות $H_1: \sigma_1^2 > \sigma_2^2$

רק אחרי שגילינו שאנו במקרה ג,

נערוך מבחן T שיבדוק את הטענה לגבי ההבדל בשכר:

(10) ב. בדקו את הטענה שתוחלת השכר של גברים גבוהה מזו של נשים בפחות מ 4000 ₪

$$H_0: \mu_1 - \mu_2 = 4000$$

$$H_1: \mu_1 - \mu_2 < 4000$$

נזכיר:

מותר לעבוד עם PV מהפלט רק אם יש 0 בהשערות

(זה לא המקרה אצלנו)

נדחה H_0 כאשר:

$$t = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} < -t_{(df, 1-\alpha)}^{\text{לסט}}$$

הסטטיסטי t המחושב
במדגם

הערך הקריטי
 t מלוח

$\bar{X}$

women	Method	N	Mean	Std Dev
0		n ₁ 9298	$\bar{X}_1$ 8862.6	s ₁ 8628.0
1		n ₂ 8197	$\bar{X}_2$ 5498.9	s ₂ 5220.3

M_1
 M_2
 (e)

$t = -5.97$
לסט

הגדל t_c לוח t :
לסט

t_c (15575, 0.95)
 1.645

$\left\{ \begin{array}{l} df \quad 15575 \\ n'' \quad \text{נהפך} \\ \alpha = 0.05 \\ 1 - 0.05 = 0.95 \end{array} \right.$

נדחה H_0 כאשר :

$$t = \frac{\bar{X}_1 - \bar{X}_2 - a}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} < -t_{(df, 1-\alpha)}^{4000}$$

הסטטיסטי t המחושב
במדגם

הערך הקריטי
 t מלוח

$$-5.97 < -1.645$$

t סט קריטי

קוחין מא
נהגים ולא כמטנקי

בר"מ 5% תוחלת השכר של גברים גבוהה מזו של נשים בפחות מ 4000 ₪

(4) ג. מהו ה p value לבדיקת הטענה בסעיף הקודם

(2) למצא Pr של טענה
נעביר t סט עדיין -5.97
נציב df (מהם) 5151
נבא על 0.5% א t נצ"ק של $Pr < 0.5\%$

(4) ד. חשבו רווח סמך ברמת סמך של 95%

רווח סמך ל $\mu_1 - \mu_2$ ברמת סמך $1 - \alpha$

(P)

$$(\bar{X}_1 - \bar{X}_2) \pm t_{(df, 1-\frac{\alpha}{2})} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$\alpha = 5\% \Rightarrow 0.05$$

$$t_{(df, 1-\frac{\alpha}{2})}$$

↓

57, 1 - 0.025

57, 0.975

57, 0.975

1.96 טי קריטי

$$3155.8 \pm 1.96 \cdot 3572.32$$

רווח הסמך סביב!

(3) ה. מהי משמעות רווח הסמך שחישבתם בסעיף הקודם

ברמת סמך של 95%

ההבדל בשכר בין לארץ לנ"ל
נ"ל בין בל. בדל. 3155.0155

(4) ו. בדקו ללא חישוב את ההשערה שתוחלת השכר של נשים בישראל שונה מתוחלת השכר של נשים באירופה העומדת על 5,500 ₪

5% α

women	Method	Mean	95% CL Mean
0		8862.6	8687.2 9038.0
1		5498.9	5385.9 5612.0

השערה נ"ל

הא דא"ר $\mu = 5500$
כוח סלק דא"ר $\mu \neq 5500$

ניסוי מניב דא"ר על כוח הסלק של הישראלי: 5500

$$5385.9 \leq \mu \leq 5612.0$$

נשים

האסיה להפסיד נכנס זמזק: 5500
כוח סלק: 5500

כחומר שכל נשים ישתל זהה ל-5500
ואם כשהבדל 5% α

אם האסיה - להפסיד לחל מניב האסיה: 5500 קוא"ל

מבחן T שני מדגמים תלויים (מזווגים)

בדיקת השערות ל μ_D - הפרש תוחלות $\mu_1 - \mu_2$ במדגמים תלויים (מזווגים)

מבחן מ 2022 מועד ב תשפ"ב

בחברת "RUN" בדקו את הטענה כי נעלי הריצה החדשות משפרות את מהירות הריצה של האצנים.

לשם כך דגמה החברה 6 אצנים שרצו מרחק שווה ללא הנעליים החדשות ולמחרת רצו עם הנעליים החדשות.

להלן התוצאות (מהירות זמן הריצה בשניות)

אצן	1	2	3	4	5	6
נעלים רגילות	56	53	64	54	49	54
נעלי RUN	53	52	56	54	48	49

(12) א. בדקו את הטענה ברמת מובהקות של 1%

$$H_0: \mu_d = 0$$

$$H_1: \mu_d > 0$$

$$\bar{d} = 3$$

$$S_d = 3.033$$

טי סטטיסטי מהנוסחה הוא 2.42

טי קריטי: דרגות חופש 5, שטח 0.99 לכן קריטי 3.36

סטטיסטי קטן מקריטי: נקבל H_0

נעלי הריצה לא משפרות

בר"מ 1%

(5) ב. חשבו את ה Pvalue לבדיקת הטענה

בעזרת טי סטטיסטי מהנוסחה הוא 2.42 וד"ח 5 נביא שטחי PV:

$$2.5\% < PV < 5\%$$

(5) ג. חשבו רווח סמך ברמת סמך של 99%

תוצאות הגבולות של הרווח:

-1.99, 7.99

(3) ד. מהי משמעות רווח הסמך שחישבתם בסעיף הקודם

ברמת סמך של 99% הפער במהירות זמן הריצה הממוצע בין נעלים רגילות לנעלי RUN

הוא בין 1.99 שניות לטובת נעלים רגילות ל-7.99 שניות לטובת נעלי RUN

תרגול עצמי, ממועד ב של שנת 2024 תשפ"ד:

שאלה 2 (32 נקודות)

חוקר טוען שגברים מחזיקים יותר זוגות נעליים מבנות זוגן .
לשם כך ערך מדגם על כמות זוגות הנעליים בקרב 5 זוגות נשואים
במגדם התקבלו התוצאות הבאות

כמות נעליים גבר	כמות נעליים אישה				
8	7				
15	3				
3	5				
7	3				
5	6				

(10) א. בדקו את הטענה ברמת מובהקות של 5%

(4) ב. חשבו את ה Pvalue לבדיקת הטענה

(5) ג. חשבו רווח סמך ברמת סמך של 95%

(4) ד. מהי משמעות רווח הסמך שחישבתם בסעיף הקודם

(5) ה. בדקו את הטענה כי לגברים יש 10 זוגות נעליים יותר מאשר לנשים.

(4) ו. בדקו את ההשערה בסעיף הקודם בדרך נוספת

פלט מבחן T שני מדגמים תלויים (מזווגים):

הועלתה הטענה כי רמת הציונים באוניברסיטה גבוהה יותר מרמת הציונים בבגרות. לשם כך נלקחו נתוני ציון הבגרות והציון באוניברסיטה למדגם של 1000 סטודנטים .
להלן פלט שבעזרתו ניתן לבדוק את ההשערה (כאשר grade = ציון באוניברסיטה avg = ציון בגרות)

The SAS System					
The TTEST Procedure					
Difference: grade - avg					
N	Mean	Std Dev	Std Err	Minimum	Maximum
1000	0.7459	8.8353	0.2794	-26.3840	23.2933

Mean	95% CL Mean	Std Dev	95% CL Std Dev
0.7459	0.1977 1.2942	8.8353	8.4643 9.2405

DF	t Value	Pr > t
999	2.67	0.0077

חשוב:

H1 היא השערה חד צדדית, לכן חובה לחלק את PV ב 2 ורק אז להשוות אותה מול רמת המובהקות אלפא.

כעת נעבוד על חלק א במבחן (טענות):

יש כאן את ההיגדים הכי חשובים ונפוצים:

בדיקת השערות ורווח סמך לשוויון תוחלות (הפרש התוחלות $\mu_1 - \mu_2$) במדגמים בלתי תלויים

שונויות ידועות $\sigma_1^2, \sigma_2^2 \Leftarrow$ ידועות

[ממועד א 2023:

שאלה 1 (30 נקודות – כל סעיף 5 נקודות)

לפניהם 6 טענות. לכל טענה יש לציין נכון/לא נכון/ לא ניתן לדעת ולנמק.

תשובה ללא נימוק לא תתקבל.

א. על מנת לבדוק את הבדלי השכר בין נשים לגברים, הוצג הפלט הבא :

	שכר גברים	שכר נשים
Mean	23.50381481	15.99465741
Known Variance	400	225
Observations	1080	1080
Hypothesized Mean Difference	9.5	
z	-2.617030551	
P(Z<=z) one-tail		
z Critical one-tail	1.644853627	
P(Z<=z) two-tail		
z Critical two-tail	1.959963985	

זד סטטיסטי

נכון/ לא נכון/אי אפשר לדעת

המספר שחסר בריבוע המודגש הוא 0.0045 :

$H_0: \mu_1 - \mu_2 = 9.5$
 $H_1: \mu_1 - \mu_2 \neq 9.5$
באט גלוי, דקרה א: סינ' א אכ
יקח
לגין
נ 3 א
pr : דו' :

$$H_0: \mu_1 - \mu_2 = a$$

$$H_1: \mu_1 - \mu_2 \neq a$$

נדחה H_0 כאשר :

$$|Z| = \frac{|\bar{X}_1 - \bar{X}_2 - a|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > Z_{1-\frac{\alpha}{2}}$$

הסטטיסטי Z
המחושב במדגם

הערך הקריטי
 Z מלוח

$$Pvalue = (\alpha') = 2 \cdot (1 - \Phi(|Z_{\bar{X}_1 - \bar{X}_2}|))$$

PK

$$Z \text{ טבל'ים } \sqrt{0.7} = -2.61$$

$$2.61$$

$$0.9955 \text{ ב } Z$$

$$Pvalue = (\alpha') = 2 \cdot (1 - \Phi(|Z_{\bar{X}_1 - \bar{X}_2}|))$$

PK

$$PV = 2 \cdot (1 - 0.9955)$$

$$PV = 0.009 \quad \checkmark$$

ההיגד שגוי!

ג. ידוע כי מספר שעות ההתנדבות השנתי של העובדים בקהילה מתפלג נורמלית עם ממוצע של 40 שעות וסטיית תקן של 14. כלכלן רשת חנויות, מעוניין לבדוק את הטענה שעובדי סניפיו מתנדבים פחות מעובדים אחרים בקהילה. לשם כך הוא אסף מסניפיו מדגם של 6 עובדים והתברר שהם התנדבו 35.5 שעות בממוצע במהלך השנה האחרונה.

Pvalue לבדיקת הטענה הינו : 0.1685

נכון/לא נכון/אי אפשר לדעת

בדיקת השערות לתוחלת μ - שונות ידועה

ההיגד שגוי:

$$z = \frac{\bar{X} - a}{\frac{\sigma}{\sqrt{n}}} = \frac{35.5 - 40}{\frac{14}{\sqrt{6}}} = -0.79$$

$$P \text{ value} = 1 - \Phi(Z) = 1 - 0.7852 = 0.2148$$

$$Pv=0.2148$$

ב. נתון רווח הסמך ברמת סמך של 99% [200 , 500] מכאן ניתן להסיק כי נדחה ברמת מובהקות של 1% את ההשערה

$$H_0: \mu = 100$$

$$H_1: \mu > 100$$

נכון/ לא נכון/אי אפשר לדעת

אם ההשערות היו דו"צ

היינו אומרים ש 100 מחוץ לתחום רווח הסמך (לא בין 200 ל 500) ולכן דוחים H_0 .

בפועל H_1 חד צדדית כלומר איזור דחיית H_0 כפול.....

רחב יותר מאשר בדו"צ אז עוד יותר בטוח שנדחה את H_0 .

גם בחד צדדי נדחה את H_0 ההיגד נכון.

אם דוחים בדו"צ בטוח דוחים בחד צדדי...

ב. נתון רווח הסמך ברמת סמך של 99% [200 , 500]
מכאן ניתן להסיק כי נדחה ברמת מובהקות של 5% את ההשערה

$$H_0: \mu = 600$$

$$H_1: \mu \neq 600$$

נכון/ לא נכון/אי אפשר לדעת

600 מחוץ לרווח הסמך של 99% רמת סמך.
כעת רמת הסמך הופכת ל 95% אז רווח הסמך מתכווץ
ובכלל בטוח ש 600 מחוץ לרווח הסמך המצומצם הזה.
ההיגד נכון.

ג. שני חוקרים דגמו מאוכלוסיות שונות מדגמים בגודל זהה וקיבלו ממוצע מדגם זהה.
שניהם בנו רווח סמך לפרמטר μ באותה רמת מובהקות. אצל חוקר א' סטיית התקן של האוכלוסייה גדולה
מזו של האוכלוסייה אצל חוקר ב'. אורך רווח הסמך של חוקר א' יהיה גדול יותר מזה של חוקר ב':
נכון/ לא נכון/אי אפשר לדעת

הסבר:

רווח סמך ל μ ברמת סמך $1 - \alpha$ (כאשר σ ידועה)

$$\bar{X} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

3 גורמים משפיעים על אורך רווח הסמך: סג"ר.
סטיית תקן, גודל המדגם, רמת מובהקות

אם סטיית התקן גדולה: אורך הרווח גדל
אם גודל המדגם גדל: אורך הרווח קטן
אם רמת המובהקות גדלה: אורך הרווח קטן

לשני החוקרים אותו גודל מדגם ואותה רמת מובהקות.
לחוקר א סטיית תקן גדולה יותר לכן אצלו אורך הרווח ארוך וגדול יותר!

ג. שני חוקרים דגמו מאותה אוכלוסייה, שני מדגמים שונים בגודל זהה.

שניהם בנו רווח סמך עם שונות ידועה לפרמטר μ באותה רמת מובהקות.

רווחי הסמך של שני החוקרים יהיו בוודאות זהים : נכון/ לא נכון/ אי אפשר לדעת

ההיגד שגוי.

אורך רווח הסמך מושפע מסטיית התקן, גודל המדגם ורמת המובהקות.

לכן האורך של שני רווחי הסמך יהיה זהה.

עם זאת, גבולות רווח הסמך יהיו שונים כי המדגמים שונים ולכן ממוצע המדגם שונה בין החוקרים.

ממוצע המדגם הוא חלק מנוסחת רווח סמך.

דוגמה לשני רווחי סמך שונים אך עם אורך זהה:

4,9 5=L אורך 5

11,16 5=L אורך 5

ד. חוקרת בדקה השערה חד צדדית עם סטיית תקן ידועה ברמת מובהקות של 5% ולא דחתה את השערת האפס החוקרת הגדילה את המדגם והממוצע לא השתנה. מכאן ניתן להסיק בוודאות שלאחר הגדלת המדגם החוקרת תדחה את השערת האפס באותה רמת מובהקות **נכון/ לא נכון/אי אפשר לדעת**

היא קיבלה את H_0 כלומר PV גדול מרמת מובהקות 5%

כעת N עלה אז PV יקטן אך לא יודעים אם עדיין ה PV החדש והמוקטן

גדול מר"מ 5% ואז שוב נקבל H_0 .

אולי ה PV החדש והמוקטן הינו קטן כעת מרמת מובהקות 5%

ואז נדחה H_0 . נסמן "לא נכון".

ד. חוקרת בדקה השערה חד צדדית עם סטיית תקן ידועה ברמת מובהקות של 5% ולא דחתה את השערת האפס התברר שסטיית התקן האמיתית קטנה יותר מזו שהשתמשה החוקרת. מכאן ניתן להסיק בוודאות שלאחר הצבת סטיית התקן הנכונה החוקרת תדחה את השערת האפס באותה רמת מובהקות **נכון/ לא נכון/אי אפשר לדעת**

במצב המוצא קיבלו את H_0 כלומר זד סטטיסטי קטן מקריטי.

לפי נוסחת Z סטיית תקן שקטנה תגרום ל זד סטטיסטי לגדול

$$\uparrow Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

לא נדע אם כעת הזד המוגדל מצליח לעבור את הקריטי כך שנדחה H_0 או שעדיין הזד המוגדל יהיה קטן מזד קריטי כך ששוב נקבל H_0 .
נסמן לא נכון.

ו. במדגם של 100 תצפיות עם שונות ידועה נבנה רווח הסמך לתוחלת השכר במשק: [8000, 10000] ₪
גודל המדגם הדרוש על מנת לבנות רווח סמך באותם נתונים עם סטייה מרבית של 250 ₪ הוא: $n=6400$
נכון/ לא נכון/ אי אפשר לדעת

פתרון רועי:

מציאת גודל המדגם n

$$n \geq \frac{\left(Z_{1-\frac{\alpha}{2}} \cdot \sigma \right)^2}{\left(\frac{L}{2} \right)^2}$$

בעזרת מחשבון ניחלץ שהנעלם סטיית התקן היא 5102.04

נציב אותה כאשר הפעם הנעלם הוא גודל המדגם N והמחשבון יראה כי גודל המדגם הדרוש הוא 1600 אנשים. ההיגד שגוי.

ו. הנהלת ערוץ טלוויזיה מעוניינת לבנות רווח בר סמך לתוחלת מספר הצופים בסדרה חדשה ברמת ביטחון 99% עם סטייה מרבית של 10 צופים. ידוע כי סטיית התקן באוכלוסייה הינה 40 צופים.
גודל המדגם הדרוש על מנת לבנות את רווח הסמך הוא 107 צופים : נכון/ לא נכון/ אי אפשר לדעת

$$n \geq \left(\frac{Z_{1-\frac{\alpha}{2}} \cdot \sigma}{\frac{L}{2}} \right)^2 = \left(\frac{2.579 \cdot 40}{10} \right)^2 = 106.17 \rightarrow n = 107$$

הטענה נכונה

ה. רמת הסמך בפלט הנתון היא 95%

נכון/ לא נכון/אי אפשר לדעת

Variable: child (child)					
N	Mean	Std Dev	Std Err	Minimum	Maximum
6901	2.3718	1.4819	0.0178	1.0000	12.0000

Mean	CL Mean		Std Dev	CL Std Dev	
2.3718	2.3369	2.4068	1.4819	1.4576	1.5070

DF	t Value	Pr > t
6900	-35.21	<.0001

אורך רווח הסמך

$$L = L_2 - L_1 = 2 \cdot t_{(n-1, 1-\frac{\alpha}{2})} \cdot S_{\bar{x}}$$

$$L = 2.4068 - 2.3369 = 2 \cdot t_{(6900, 1-\frac{\alpha}{2})} \cdot 0.0178 \rightarrow t_{(6900, 1-\frac{\alpha}{2})} = 1.96 \rightarrow \alpha = 0.05$$

אפשר גם להציב $t_{(6900, 0.975)} = 1.96$ ולקבל את הרווח סמך המופיע בפלט

Variable: child (child)					
N	Mean	Std Dev	Std Err	Minimum	Maximum
6901	2.3718	1.4819	0.0178	1.0000	12.0000

Mean	95% CL Mean	Std Dev	95% CL Std Dev
2.3718	2.3369 2.4068	1.4819	1.4576 1.5070

DF	t Value	Pr > t
6900	-35.21	<.0001

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} =$$

$$\frac{2.3718 - \mu_0}{0.0178} = -35.21$$

לאחר שהשוונו מול טי סטטיסטי נחלץ ש M הוא 3
טענה נכונה

מבוא לסטטיסטיקה ב' - תשפ"ד
בחינה - סמסטר ב' - מועד ב' - 8.7.2024

א. על מנת לבדוק את הטענה שהשכר לשעה ב \$ של עובדים בעלי דירה גבוה מהשכר לשעה ב \$ של עובדים שאין בבעלותם דירה, הוצג הפלט הבא :

F-Test Two-Sample for Variances		
	יש דירה	אין דירה
Mean	18	13
Variance	172	109
Observations	6169	2848
df	6168	2847
F		
P(F<=f) one-tail		
F Critical one-tail		

על סמך התוצאות שהתקבלו בפלט ניתן להסיק כי שונות השכר לשעה בין שני סוגי העובדים זהה

נכון/ לא נכון/ אי אפשר לדעת

א. על מנת לבדוק את הטענה שהשכר לשעה ב \$ של עובדים בעלי דירה גבוה מהשכר לשעה ב \$ של עובדים שאין בדירותם דירה, הוצג הפלט הבא:

F-Test Two-Sample for Variances	1	2
	יש דירה	אין דירה
Mean	18	13
Variance	172	109
Observations	6169	2848
df	6168	2847
F		
P(F<=f) one-tail		
F Critical one-tail		

על סמך התוצאות שהתקבלו בפלט ניתן להסיק כי שונות השכר לשעה בין שני סוגי העובדים זהה

נכון/ לא נכון/ אי אפשר לדעת

למק קל הניסוח:

בדיקת השערות לשוויון שונות ($\sigma_1^2 = \sigma_2^2$) במדגמים בלתי תלויים

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 > \sigma_2^2$$

נדחה H_0 כאשר:

$$F = \frac{S_1^2}{S_2^2} = \frac{\frac{\sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2}{n_1 - 1}}{\frac{\sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2}{n_2 - 1}} > F(n_1 - 1, n_2 - 1, 1 - \alpha)$$

הסטטיסטי F המחושב במדגם הערך הקריטי F מלוח

סגור:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 > \sigma_2^2$$

חישוב F

$$F = \frac{S_1^2}{S_2^2} = \frac{172}{109} = 1.57$$

נקודה קריטית על ידי התפלגות F

$$F_c = 1$$

$F(n_1 - 1, n_2 - 1, 1 - \alpha)$
 6168, 2847, 0.95
 ד"ה נ"ה (ד"ה נ"ה)
 ד"ה נ"ה (ד"ה נ"ה)
 הערך הקריטי
 F מלוח

$$F_{\text{calc}} > F_c$$

1.57 > 1
 ק"ק

ס"ס דמיון
 מ

נקדה מ: סיונו שיו
 ה 5% = 0.05

p	n	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	120	∞
.95	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.79	2.75	2.71
.975		7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.56	3.45	3.39	3.33
.99		10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.65	4.48	4.40	4.31
.995		13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23	6.03	5.83	5.62	5.41	5.30	5.19
.95	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.70	2.62	2.58	2.54
.975		6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.31	3.20	3.14	3.08
.99		10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.25	4.08	4.00	3.91
.995		12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66	5.47	5.27	5.07	4.86	4.75	4.64
.95	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.47	2.38	2.34	2.30
.975		6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	2.96	2.85	2.79	2.72
.99		9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.70	3.54	3.45	3.36
.995		11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91	4.72	4.53	4.33	4.12	4.01	3.90
.95	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.16	2.11	2.07
.975		6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.64	2.52	2.46	2.40
.99		8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.21	3.05	2.96	2.87
.995		10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25	4.07	3.88	3.69	3.48	3.37	3.26
.95	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.95	1.90	1.84
.975		5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.35	2.22	2.16	2.09
.99		8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.78	2.61	2.52	2.42
.995		9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.68	3.50	3.32	3.12	2.92	2.81	2.69
.95	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.74	1.68	1.62
.975		5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.07	1.94	1.87	1.79
.99		7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.39	2.21	2.11	2.01
.995		9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.18	3.01	2.82	2.63	2.42	2.30	2.18
.95	60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.65	1.53	1.47	1.39
.975		5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.82	1.67	1.58	1.48
.99		7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.03	1.84	1.73	1.60
.995		8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.74	2.57	2.39	2.19	1.96	1.83	1.69
.95	120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.55	1.43	1.35	1.25
.975		5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.69	1.53	1.43	1.31
.99		6.85	4.79	3.95	3.48	3.17	2.95	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.86	1.66	1.53	1.38
.995		8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54	2.37	2.19	1.98	1.75	1.61	1.43
.95	∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.76	1.67	1.57	1.46	1.32	1.22	1.00
.975		5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.57	1.39	1.27	1.00
.99		6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.70	1.47	1.32	1.00
.995		7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36	2.19	2.00	1.79	1.53	1.36	1.00

ה. שני חוקרים בנו על סמך אותו מדגם של 15 תצפיות רווח סמך לפרמטר μ באותה רמת סמך.

לחוקר א' היה מידע על סטיית התקן של האוכלוסייה ואילו החוקר ב' אמד את סטיית התקן ע"י המדגם, אך התוצאה הייתה זהה. מכאן ניתן להסיק שרווח הסמך של חוקר ב' יהיה ארוך יותר מרווח הסמך של חוקר א' **נכון/ לא נכון/ אי אפשר לדעת**

גודל המדגם N הוא קטן מ 120

לכן ערך T גדול מערך Z

לאותה רמת מובהקות.

$$n < 120 \quad \text{אז}$$

$$Z < t$$

סטיית התקן באוכלוסייה G זהה לסטיית התקן במדגם S: זה נתון.

גודל המדגם N זהה לשני החוקרים.

רמת הסמך זהה לשני החוקרים.

לכן חוקר ב אשר עובד עם נוסחת רווח סמך של T

יקבל אורך גבוה מחוקר א שעובד עם נוסחת רווח סמך Z

וזאת כאמור כי $T > Z$.

הטענה נכונה.

אגב, לידיעה:

אם $N < 120$ אז ערך T שווה לערך Z

ד. שתי חוקרות בדקו באותו מדגם של 200 תצפיות ובאותה רמת מובהקות את ההשערה

$$H_0: \mu = 4,000$$

$$H_1: \mu > 4,000$$

חוקרת א' בדקה את ההשערה בהנחה שסטיית התקן ידועה וחוקרת ב' אמדה את סטיית התקן במדגם

התברר כי סטיית התקן שנאמדה הייתה נמוכה יותר מסטיית התקן הידועה

אם חוקרת א' תדחה את השערת האפס גם חוקרת ב' תדחה את השערת האפס

נכון / לא נכון/אי אפשר לדעת

חוקרת א עובדת עם סטיית תקן אוכלוסיה כלומר עם Z

חוקרת א דחתה H0 סימן שזד סטטיסטי אצלה גדול מ Z קריטי שלה

לפי הנתונים- חוקרת ב עובדת עם סטיית תקן מדגם S שהיא נמוכה מסטיית התקן באוכלוסיה.

כלומר טי סטטיסטי של חוקרת ב גבוה מזד סטטיסטי חוקרת א, הסבר:

$$t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}} > Z$$

N>120 לכן T קריטי = Z קריטי

חוקרת א הצליחה לעבור את הזד הקריטי בעזרת הזד הסטטיסטי שלה.
לחוקרת ב יש טי סטטיסטי שהוא גדול מהזד הסטטיסטי.
לכן גם בטוח שחוקרת ב תעקוף את אותו קריטי....
לכן חוקרת ב גם תדחה H_0 הטענה נכונה.