

The present report presents statistical analyses to test diverse hypotheses based in a survey done in some years ago at Yongin area. The observers noted whether a vehicle stops on red light at the exact stop line or not. They also noted the types and brands of vehicles and the biographical information of the drivers. The different hypothesis here presented are evaluated using the Chi-square (X^2) statistical test.

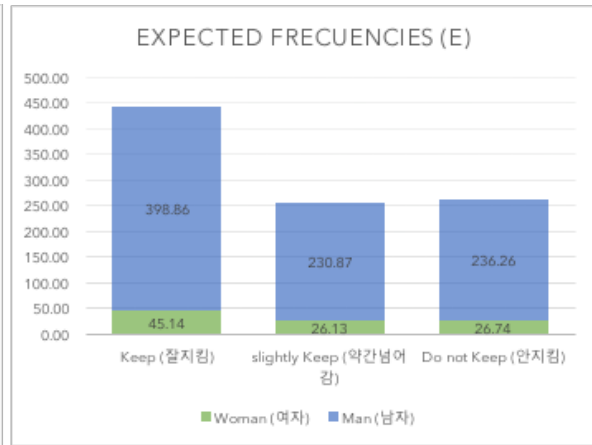
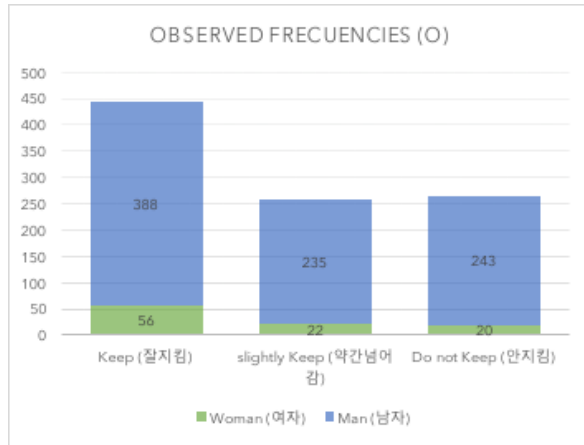
Hypothesis 1: Male drivers are more likely to come to a full stop at the stop line than female drivers.

Observed Frequencies (O)	Woman (여자)	Man (남자)	Total man and Woman
Keep (잘지킴)	56	388	444
slightly Keep (약간넘어감)	22	235	257
Do not Keep (안지킴)	20	243	263
Total Cases	98	866	964

Observed Frequencies (%)	Woman (여자)	Man (남자)
Keep (잘지킴)	57%	45%
slightly Keep (약간넘어감)	22%	27%
Do not Keep (안지킴)	20%	28%

Expected Frequencies (E)	Woman (여자)	Man (남자)	Total man and Woman
Keep (잘지킴)	45.14	398.86	444
slightly Keep (약간넘어감)	26.13	230.87	257
Do not Keep (안지킴)	26.74	236.26	263
Total Cases	98	866	964

Chi-square (X^2)	Woman (여자)	Man (남자)
Keep (잘지킴)	2.61	0.30
slightly Keep (약간넘어감)	0.65	0.07
Do not Keep (안지킴)	1.70	0.19
Total (X^2)		5.525



Chi-Square(X²) 5.525
Probability (P) 0.063
Degrees of Freedom (df) 2

Results.

Since the P-value (0.063) is greater than the significance level (0.05), we accept the null hypothesis. Thus, we conclude that there is a relationship between gender and keeping the stop line. Also, the test results show that the differences are likely to change.

Hypothesis 2: The older drivers are more likely to observe the stop line than the younger drivers.

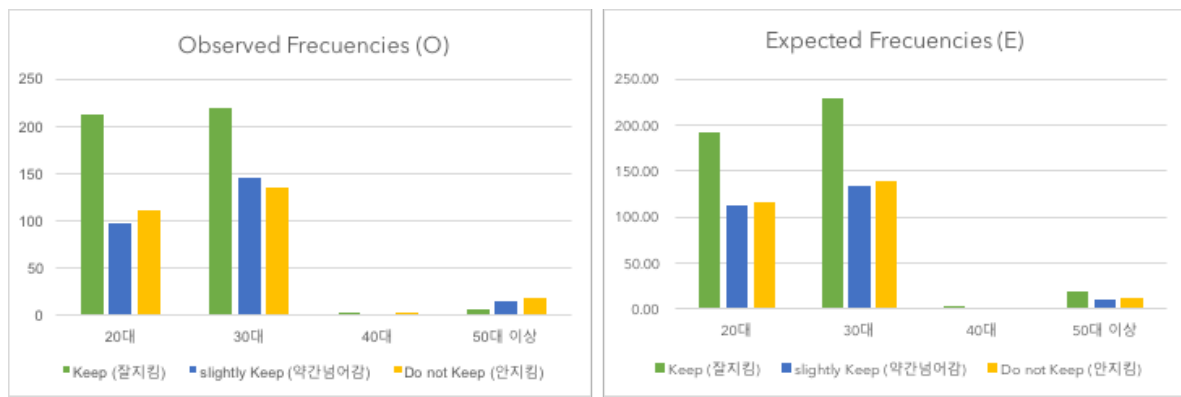
Observed Frequencies (O)	20 대	30 대	40 대	50 대 이상	Total
Keep (잘지킴)	213	220	3	6	442
slightly Keep (약간넘어감)	98	146	0	15	259
Do not Keep (안지킴)	111	135	2	19	267
Total Cases	422	501	5	40	968

Observed Frequencies (%)	20 대	30 대	40 대	50 대 이상
Keep (잘지킴)	50%	44%	60%	15%
slightly Keep (약간넘어감)	23%	29%	0%	38%
Do not Keep (안지킴)	26%	27%	40%	48%

Expected Frequencies (E)	20 대	30 대	40 대	50 대 이상	Total
Keep (잘지킴)	192.69	228.76	2.28	18.26	442
slightly Keep (약간넘어감)	112.91	134.05	1.34	10.70	259
Do not Keep (안지킴)	116.40	138.19	1.38	11.03	267
Total Cases	422	501	5	40	968

<i>Chi-square (X2)</i>	20 대	30 대	40 대	50 대 이상
Keep (잘지킴)	2.14	0.34	0.23	8.24
slightly Keep (약간넘어감)	1.97	1.07	1.34	1.73
Do not Keep (안지킴)	0.25	0.07	0.28	5.75
Total (X2)	23.39			

Chi-Square(X2) 23.39
Probability (P) 0.00068
Degrees of Freedom (df) 6



Results.

Since the P-value (0.0006) is less than the significance level (0.05), we cannot accept the null hypothesis. Thus, we conclude that there is a relationship between age and keeping the stop line.

Hypothesis 3: Observing the stop line depends on the color of vehicles.

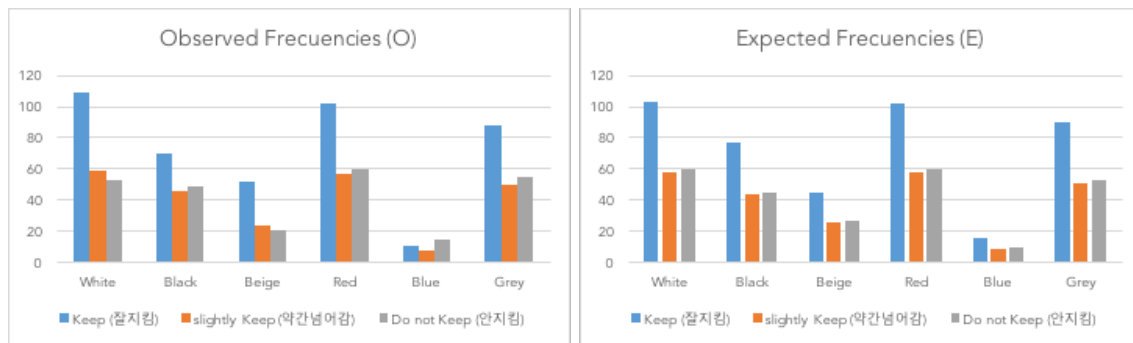
<i>Observed Frequencies (O)</i>	White	Black	Beige	Red	Blue	Grey	Total
Keep (잘지킴)	109	70	52	102	11	88	432
slightly Keep (약간넘어감)	59	46	24	57	8	50	244
Do not Keep (안지킴)	53	49	21	60	15	55	253
Total Cases	221	165	97	219	34	193	929

<i>Observed Frequencies (%)</i>	White	Black	Beige	Red	Blue	Grey
Keep (잘지킴)	49%	42%	54%	47%	32%	46%
slightly Keep (약간넘어감)	27%	28%	25%	26%	24%	26%
Do not Keep (안지킴)	24%	30%	22%	27%	44%	28%

Expected Frequencies (E)	White	Black	Beige	Red	Blue	Grey	Total
Keep (잘지킴)	102.77	76.73	45.11	101.84	15.81	89.75	432
slightly Keep (약간넘어감)	58.05	43.34	25.48	57.52	8.93	50.69	244
Do not Keep (안지킴)	60.19	44.94	26.42	59.64	9.26	52.56	253
Total Cases	221	165	97	219	34	193	929

Chi-square (X2)	White	Black	Beige	Red	Blue	Grey	
Keep (잘지킴)	0.38	0.59	1.05	0.00	1.46	0.03	
slightly Keep (약간넘어감)	0.02	0.16	0.09	0.00	0.10	0.01	
Do not Keep (안지킴)	0.86	0.37	1.11	0.00	3.56	0.11	
Total (X2)							9.91

Chi-Square(X2) 9.91
Probability (P) 0.45
Degrees of Freedom (df) 10



Results.

Since the P-value (0.45) is greater than the significance level (0.05), we accept the null hypothesis. Thus, we conclude that there is a relationship between where the vehicles come from and keeping the stop line. Also, the test results show that the differences are likely to change.

Hypothesis 4: There are differences in observing stop line among these four groups: older males, younger males, older females, and younger females.

Observed Frequencies (O)	20 대	30 대	40 대	50 대 이상	Total
Keep (잘지킴)	213	220	3	6	442
slightly Keep (약간넘어감)	98	146	0	15	259
Do not Keep (안지킴)	111	135	2	19	267
Total Cases	422	501	5	40	968

<i>Observed Frequencies (O)</i>	Woman (여자)	Man (남자)	Total
Keep (잘지킴)	56	388	444
slightly Keep (약간넘어감)	22	235	257
Do not Keep (안지킴)	20	243	263
Total Cases	98	866	964

However, according to the hypothesis we need to divided the data in two categories, young and old Males and females respectively. Young are classified to be in their 20-30's and old in their 40-50's.

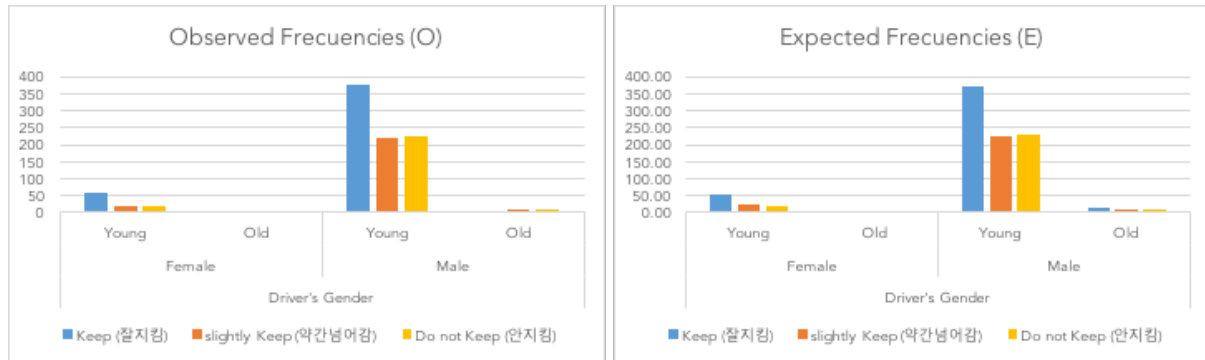
<i>Observed Frequencies (O)</i>	Driver's Gender					
	Female			Male		
	Young	Old	Female Total	Young	Old	Male Total
Keep (잘지킴)	56	0	56	375	7	382
slightly Keep (약간넘어감)	21	1	22	222	10	232
Do not Keep (안지킴)	20	0	20	224	11	235
Total Cases	97	1	98	821	28	849

<i>Observed Frequencies %</i>	Driver's Gender			
	Female		Male	
	Young	Old	Young	Old
Keep (잘지킴)	58%	0%	46%	25%
slightly Keep (약간넘어감)	22%	100%	27%	36%
Do not Keep (안지킴)	21%	0%	27%	39%

<i>Expected Frequencies (E)</i>	Driver's Gender					
	Female			Male		
	Young	Old	Female Total	Young	Old	Male Total
Keep (잘지킴)	55.43	0.57	56.00	369.40	12.60	382.00
slightly Keep (약간넘어감)	21.78	0.22	22.00	224.35	7.65	232.00
Do not Keep (안지킴)	19.80	0.20	20.00	227.25	7.75	235.00
Total Cases	97	1	98	821	28	849

<i>Chi-square (X2)</i>	Driver's Gender			
	Female		Male	
	Young	Old	Young	Old
Keep (잘지킴)	0.01	0.57	0.08	2.49
slightly Keep (약간넘어감)	0.03	2.68	0.02	0.72
Do not Keep (안지킴)	0.00	0.20	0.05	1.36
Total (X2)	8.22			

Chi-Square(X²) 8.22
 Probability (P) 0.065
 Degrees of Freedom (df) 6



Results.

Since the P-value (0.065) is greater than the significance level (0.05), we accept the null hypothesis. Thus, we conclude that there is a relationship between gender/Age and keeping the stop line. Also, the test results show that the differences are likely to change.

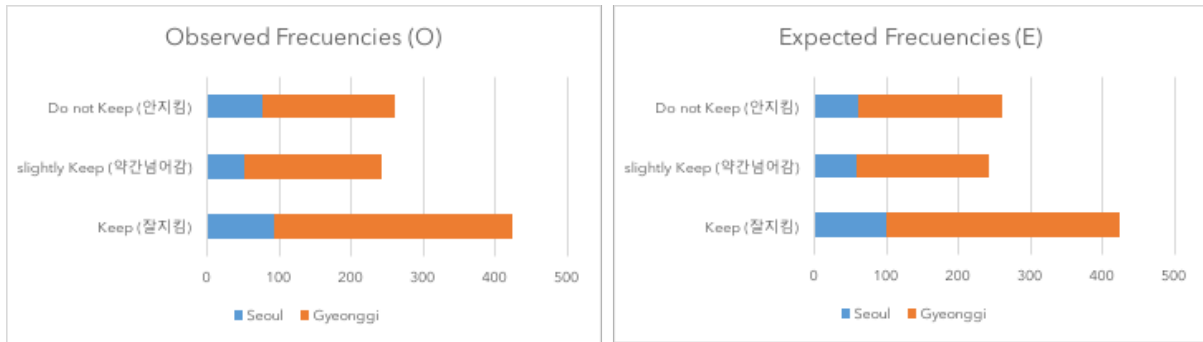
Hypothesis 5: The vehicles from Seoul are more likely to observe the stop line than those of Kyonggi Province.

Observed Frequencies (O)	Seoul	Gyeonggi	Total
Keep (잘지킴)	92	332	424
slightly Keep (약간넘어감)	51	192	243
Do not Keep (안지킴)	76	185	261
Total Cases	219	709	928

Observed Frequencies (%)	Seoul	Gyeonggi	
Keep (잘지킴)	42%	47%	
slightly Keep (약간넘어감)	23%	27%	
Do not Keep (안지킴)	35%	26%	
Expected Frequencies (E)	Seoul	Gyeonggi	Total
Keep (잘지킴)	100.06	323.94	424
slightly Keep (약간넘어감)	57.35	185.65	243
Do not Keep (안지킴)	61.59	199.41	261
Total Cases	219	709	928

Chi-square (X2)	Seoul	Gyeonggi	
Keep (잘지킴)	0.65	0.20	
slightly Keep (약간넘어감)	0.70	0.22	
Do not Keep (안지킴)	3.37	1.04	
Total (X2)		6.18	

Chi-Square(X2) 6.18
Probability (P) 0.05
Degrees of Freedom (df) 2



Results.

Since the P-value (0.05) is same than the significance level (0.05), we fail to reject the null hypothesis. Thus, we conclude that there is a relationship between where are the vehicles come from and keeping the stop line.

Hypothesis 6: Domestic car drivers are more likely to observe the stop line than foreign car drivers.

Observed Frequencies (O)	Hyundai	Daewoo	Kia	Ssangyong	Foreign Brand	Total
Keep (잘지킴)	183	105	81	4	24	397
slightly Keep (약간넘어감)	85	71	50	1	15	222
Do not Keep (안지킴)	110	39	61	4	12	226
Total Cases	378	215	192	9	51	845

However, according to the hypothesis we need to divide the data in two categories. Domestic and Foreign brands of cars.

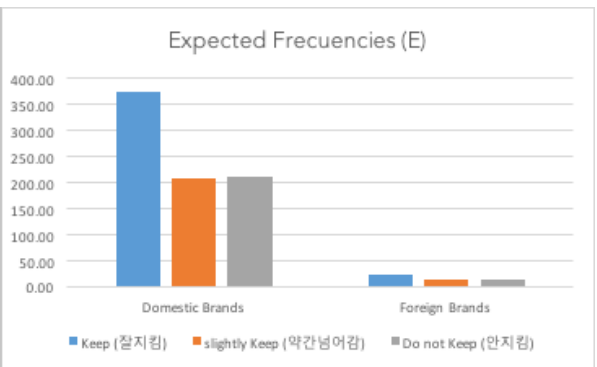
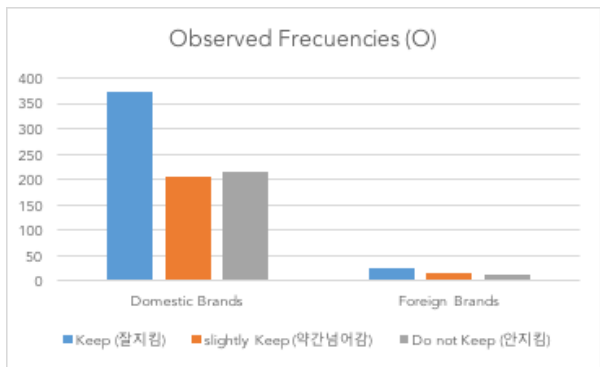
<i>Observed Frequencies (O)</i>	Domestic Brands	Foreign Brands	Total
Keep (잘지킴)	373	24	397
slightly Keep (약간넘어감)	207	15	222
Do not Keep (안지킴)	214	12	226
Total Cases	794	51	845

<i>Observed Frequencies (%)</i>	Domestic Brands	Foreign Brands
Keep (잘지킴)	47%	47%
slightly Keep (약간넘어감)	26%	29%
Do not Keep (안지킴)	27%	24%

<i>Expected Frequencies (E)</i>	Domestic Brands	Foreign Brands	Total
Keep (잘지킴)	373.04	23.96	397
slightly Keep (약간넘어감)	208.60	13.40	222
Do not Keep (안지킴)	212.36	13.64	226
Total Cases	794	51	845

<i>Chi-square (X2)</i>	Domestic Brands	Foreign Brands
Keep (잘지킴)	0.00	0.00
slightly Keep (약간넘어감)	0.01	0.19
Do not Keep (안지킴)	0.01	0.20
Total (X2)		0.41

Chi-Square(X2) 0.41
Probability (P) 0.81
Degrees of Freedom (df) 2



Results.

Since the P-value (0.04) is less than the significance level (0.05), we cannot accept the null hypothesis. Thus, we conclude that there is a relationship between car Brand and keeping the stop line.

(Extra) Hypothesis 7: Observing the stop line depends on the number of passengers in the vehicle.

<i>Observed Frequencies (O)</i>	1 명	2 명	3 명 이상	Total
Keep (잘지킴)	268	178	3	449
slightly Keep (약간넘어감)	137	115	9	261
Do not Keep (안지킴)	135	125	11	271
Total Cases	540	418	23	981

<i>Observed Frequencies (%)</i>	1 명	2 명	3 명 이상	
Keep (잘지킴)	50%	43%	13%	
slightly Keep (약간넘어감)	25%	28%	39%	
Do not Keep (안지킴)	25%	30%	48%	

<i>Expected Frequencies (E)</i>	1 명	2 명	3 명 이상	Total
Keep (잘지킴)	247.16	191.32	10.53	449
slightly Keep (약간넘어감)	143.67	111.21	6.12	261
Do not Keep (안지킴)	149.17	115.47	6.35	271
Total Cases	540	418	23	981

<i>Chi-square (X2)</i>	1 명	2 명	3 명 이상	
Keep (잘지킴)	1.76	0.93	5.38	
slightly Keep (약간넘어감)	0.31	0.13	1.36	
Do not Keep (안지킴)	1.35	0.79	3.40	
Total (X2)			15.39	

Chi-Square(X2) 15.39
Probability (P) 0.017
Degrees of Freedom (df) 4

Results.

Since the P-value (0.017) is less than the significance level (0.05), we cannot accept the null hypothesis. Thus, we conclude that there is a relationship between number of passengers and keeping the stop line.

