Sample: Statistics- Hypothesis Testing

The study presents findings from interviews of 52 divorced individuals who received the Prevention and Relationship Enhancement Program (PREP) while engaged to be married

1. What test of significance did the researchers use?

Independent t test and calculation of kappa statistic.

Research question of t test
The approach of reporting a result requires whether or not result is statistically significant or not.

Independent t test is used to compare the 2 sample**s.**
First sample is of 52 divorced individuals who received the PREP while engaged to be married.
Second sample consists of individuals who did not participated in any kind of course. The research question here is whether these 2 groups have different prpoperties or not.

2. Findings of t test
Null and alternative hypotheses for the independent t-test
Null Hypothesis (population means are equal)

$$H_0: u_1 = u_2$$

Alternate Hypothesis (population means are not equal)

$$H_A: u_1 \neq u_2$$

Meaning of population mean in our context of study
Both of the groups can be compared based on a number of attributes, ones which are used in this study are:
Age at marriage, ethnicity, personal income and relationship adjustment at the premarital assesement.

Polulation mean refers to the mean of both the samples across different attributes.

Quoting the exact statement written in the paper.
(Reasons for Divorce and Recollections of Premarital Intervention: Implications for Improving Relationship Education
Shelby B. Scott, Galena K. Rhoades, Scott M. Stanley, Elizabeth S. Allen, and Howard J. Markman)

There were no significant differences between divorced individuals who participated in this study compared to divorced individuals who did not participate across age at marriage, ethnicity, personal income, or relationship adjustment at the premarital assessment.
Means there was no significance difference in income/ethnicity, personal income/relationship adjustment/age at marriage between people who took PREP course and who didn't

(explanation:

2 column were prepared for say personal income, one contains the personal income of people who took PREP course and one for people who didn't took. And then t statistic is calculated. Similarly for other attributes mentioned.)

Finding with regard to t statistic

The significance level i.e value of α for t test is 5%. Interpretation of this is:

There is only 5% chance that there exists a significant difference between the two test groups and we were not able to detect it.

p-value obtained for this test in the dataset was greater than 0.05 hence, Null Hypothesis was accepted.

Grounded Theory and its implementation

Derieved from (Creswell, 2006; Strauss & Corbin, 1998)

Qualitative research produces findings that are not arrived at by statistical procedures. It can refer to research about persons' lives, lived experiences, behaviors, emotions, and feelings as well as about organizational functioning, social movements, cultural phenomena, and interactions between nations.

It mean theory that was derived from data, systematically gathered and analyzed through the research process. The researcher begins with an area of study and allows the theory to emerge from the data. Theory derived from data is more likely to resemble the "reality" than is theory derived by putting together a series of concepts based on experience or solely through speculation.

Test statistic in grounded theory

(explanation:

		Rater 1		
		Yes	No	
Rater 2	Yes	X_1	Y_1	rm^1
	No	X_2	Y_2	rm^2
		cm^1	cm^2	

Now X_1 is the number of questions to which both the rater 1 and rater 2 has answered yes (both agreed) similarly Y_2 is the questions to which both the rater 1 and rater 2 has answered no (again both agreed)

X_2 and Y_1 represents the number of questions to which their answers differed.

Pr(a)= Actual observed aggrement = $(X_1 + Y_2)/(X_1+X_2+Y_1+Y_2)$

Pr(e) = represents the chance aggrement = $((cm^1*rm^1)/n^2+ (cm^2*rm^2)/n^2$jjjj

Kappa Value = $(Pr(a)-Pr(e))/(1-Pr(e))$

cm^1 = column 1 marginal =(X_1+X_2)
cm^2 = column 2 marginal =(Y_1+Y_2)
rm^1 = row 1 marginal = $(X_1 + Y_1)$
rm^2 =row 2 marginal = $(X_2 +Y_2)$
n represents the number of observations (not the number of raters)

Exact lines quoted from paper
(Reasons for Divorce and Recollections of Premarital Intervention: Implications for Improving Relationship Education
Shelby B. Scott, Galena K. Rhoades, Scott M. Stanley, Elizabeth S. Allen, and Howard J. Markman)

Research question of Cohen Kappa

Using items from a previous survey on reasons for divorce (C. A. Johnson et al., 2001) participants were asked to indicate whether or not each item on a list of common problems in relationships was a "major contributor to their divorce" ("yes" or "no"). These items included lack of commitment, infidelity/extra-marital affairs, too much arguing or conflict, substance abuse, domestic violence, economic hardship, lack of support from family members, marrying too young, little or no premarital education, and religious differences. The question is here is weather there is agreement between the partners (rate 1 and rater 2) on the reasons of divorce.

Findings of Cohen Kappa

There was satisfactory evidence found that there is significant amount of agreement between both of the raters regarding the reasons of divorce.

Findings with regard to Cohen Kappa statistic

The average Cohen's Kappa was .71 (SD = .28) and the median was 0.80.
Kappa = (observed accuracy - expected accuracy)/(1 - expected accuracy)
It is a statistic which measures inter-rater agreement for qualitative (categorical) items
Interpreting Kappa Value

Kappa≤ 0 : No aggrement
0.01-0.20 : None to slight aggrement
0.21-0.40 : Fair aggrement
0.41-0.60 :Moderateaggrement
0.61-0.80 : Satisfactory aggrement
0.80-1.00 : Almost perfect aggrement

Also, the descriptive statistical results are summarised the table which contains reasons of divorce.

2. Which test of significance did the researchers use?

In this study relationship between Age at marriage and divorce rate is studied.

Cox proportional hazards models is applied to study this relationship. Hazard Ratio is calculated as test statistic known as the test ratio

The probability of the divorce is called the hazard. The hazard is modeled as:

$$H(t) = H_0(t) \text{ x } \exp(b_1 * Age)$$

$H_0(t)$ is the value of baseline hazard at time t

The Ratio $H(t)/H_0(t)$ is known as Hazard Ratio

Findings of Hazard Model
A zero-order model with a linear term for age at marriage shows a 7 percent reduction in the odds of divorce for each year of age
Linear model doesn't explains the effect of age at marriage. To determine the exact form. Piecewise linear spline model is constructed and the best fitting model has 2 piece spline with a knot at the age of 28.
Findings of the model:

- Youthful marriage is a strong predictor of divorce
- Each age of age prior to 28 reduces the odds of divorce by 11%
- Past 28, the odds of divorce increases by 5% per year
- The change in slope is statistically signifcant

Findings regarding Hazrad Ratio

This model conforms that the relationship between the considered variables is definitely not linear. Exact relationship is shown below.

Relative Odds of Divorce by Age at Marriage

Source: NSFG 2006-2010
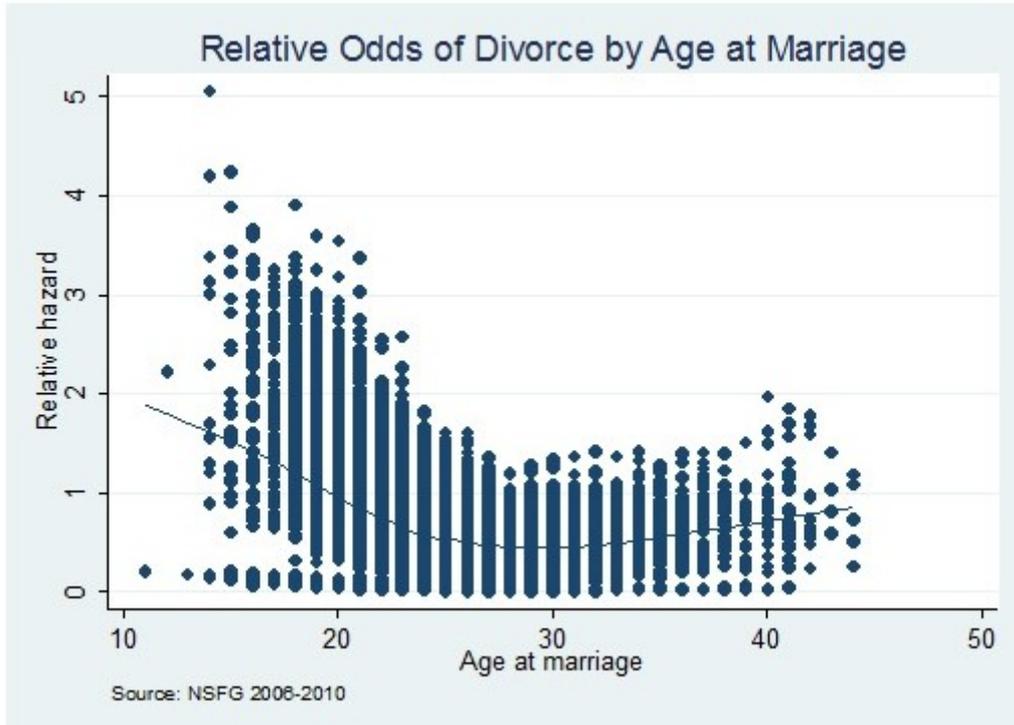
Divorces per 1,000 married people: 2012

	Age at marriage				
Years married	15-19	20-24	25-29	30-34	35-39
1-5	43	27	19	17	23
5-9	55	33	27	26	26
10-14	34	30	20	22	26
15-24	27	20	18	14	16
25-34	17	12	12	8	8

Source: Philip N. Cohen calculations from American Community Survey