One Sample & Two independent samples

van der Horst, K., Ferrage, A., & Rytz, A. (2014). Involving children in meal preparation. Effects on food intake. Appetite, 79, 18–24. https://doi.org/10.1016/j.appet.2014.03.030

Introduction

Van der Horst et al. (2014) studied the effect of children's involvement in meal preparation on their food and vegetable intake. A between-subject experiment was conducted with 47 children aged 6 to 10 years. In condition 1 (n= 25), children prepared a lunch meal with the assistance of a parent. The parent prepared the meal alone in condition 2 (n = 22). Only the caloric intake (total energy in kcal) per meal will be analysed for this homework. Random variable Y – caloric intake (kcal) per meal of children in condition 1. Random variable X - caloric intake (kcal) per meal of children in condition 2

Sample Characteristics

The skim function was used for both conditions to obtain some statistics of interest.

mean <dbl>	sd <dbl>	p0 <dbl>	p25 <dbl>	p50 <dbl>	p75 <dbl>	p100 <dbl>	hist <chr>
431.3996	105.7012	249.86	368.51	428.74	477.96	635.21	

Table 1 Statistics and histogram for Condition 1

mean <dbl>	sd <dbl>	p0 <dbl>	p25 <dbl>	p50 <dbl>	p75 <dbl>	p100 <dbl>	hist <chr>
346.7991	99.50114	139.69	290.3975	361.02	422.5675	503.46	

Table 2 Statistics and histogram for Condition 2

Shapiro-Wilks tests were applied to verify if a gaussian probability distribution adjusts to both populations. For both conditions, the null hypothesis wasn't rejected for an alpha of .5, meaning there is no statistical evidence that both populations do not follow a Gaussian distribution (W_{cond1} = .955, p-value $_{cond1}$ = .319; W_{cond2} = .958, p-value $_{cond2}$ = .451) This conclusion was supported by the analysis of the Q.Q Plot's (present in script).

T Test for two independent samples

Are there differences between both conditions in terms of the mean value of caloric intake? Given that both samples (conditions) are independent and follow a gaussian distribution, the most appropriate statistical test is the T-Test for two independent samples to test if there are differences between both conditions regarding calorie intake.

The assumption of variances was verified using the var.test function. The null hypothesis of equal variances was not rejected with an alpha of .5 (F = 1.1285, p-value = .7848). Given that both populations are homoscedastic, the test statistics used are presented in figure 6.

$$T_{obs} = \frac{\bar{X} - \bar{Y} - d}{S_p\sqrt{1/n_1 + 1/n_2}} \underset{\text{sob } H_0}{\frown} t_{(n_1+n_2-2)}$$

Equation 1 Test Statistic for two homoscedastic independent samples T Test

The hypothesis to be tested is H0: $\mu_{cond2}-\mu_{cond1}=0$ vs H1:$\mu_{cond2}-\mu_{cond1}\neq 0$. For this test, the null hypothesis was rejected for an alpha of .01, which means that there is statistical evidence that both conditions differ in terms of caloric intake (t = -2.8137, df = 45, p-value = .007). A 95% confidence interval for the mean difference was also obtained: [24.042, 145. 158].

T Test for one sample

Given that there were observed significant differences between both conditions and that the mean value of calorie intake from condition 1 is higher than the one from condition 2 (negative T), another question arises: Does the mean caloric intake value of Condition one higher than the value of reference (400)? The hypothesis to be tested is H0: $\mu_{cond1} \leq 400$ vs $\mu_{cond1} > 400$. To test this, the appropriate statistical test is the T-Test for one sample, given that the population follows a Gaussian distribution. The test statistic can be seen in Formula 2.

$$T = \sqrt{n}\,\frac{\overline{Y} - \mu_0}{S} \cap t_{n-1}$$

Equation 2 Test Statistic for one sample T Test

For this test, the null hypothesis was not rejected for an alpha of .5, which means there is no statistical evidence that the caloric intake mean value of Condition one is higher than the value of reference 400 (T = 1.485, df = 24, p-value = .075). A 95% confidence interval for the mean value was also obtained: [387.768, 475.031].

T Test for paired samples

Pasman, W. J., van Baak, M. A., Jeukendrup, A. E., & de Haan, A. (1995). The effect of different dosages of caffeine on endurance performance time. *International Journal of Sports Medicine*, *16*(4), 225–230. https://doi.org/10.1055/s-2007-972996

Introduction

In order to examine the effect of different dosages of caffeine (0-5-9-13 mg.kg body weight-1) on endurance performance, Pasman et al. (1995) observed nine well-trained cyclists. Caffeine capsules were administered in random order and double-blind. One hour after capsule ingestion, subjects cycled until exhaustion at 80% Wmax. Only two dosages were considered for this homework (0 - Placebo and 13). Random variable Y – VO2max value of athletes with 0mg of caffeine. Random variable X - VO2max value of athletes with 13mg of caffeine

Sample Characteristics

The skim function was again used for both conditions to obtain some statistics of interest.

mean <dbl>	sd <dbl>	p0 <dbl>	p25 <dbl>	p50 <dbl>	p75 <dbl>	p100 <dbl>	hist <chr>
46.44	12.48826	28.34	36.05	45.2	56.55	66.38	▇▇▂...

Table 3 Statistics and histogram for Condition No caffeine (Placebo)

mean <dbl>	sd <dbl>	p0 <dbl>	p25 <dbl>	p50 <dbl>	p75 <dbl>	p100 <dbl>	hist <chr>
58.14889	15.13416	36.2	46.48	59.3	69.47	79.12	▇▂▇...

Table 4 Statistics and histogram for Condition 13 mg of caffeine

Shapiro-Wilks test was applied to verify if a gaussian probability distribution adjusts to both populations. For both conditions, the null hypothesis wasn't rejected for an alpha of .5, meaning there is no statistical evidence that both populations do not follow a Gaussian distribution ($W_{no\_caffeine}$ = .930, $\text{p-value}_{no\_caffeine}$ = .482; W_{13mg} = .965, p-value_{13mg} = .844). The analysis of the Q.Q Plot's supported this conclusion (present in the script).

T Test for paired samples

Does the endurance (VO2max) of the athletes increase when they take 13mg of caffeine? Given that the samples are paired and they both follow a gaussian distribution, the appropriate statistical test to answer this question is the T Test for paired samples with the test statistic present in equation 3

$$T = \frac{\overline{D} - \delta_0}{S_D / \sqrt{n}} \cap t_{(n-1)}$$

Table 5 Test Statistic for two paired samples. The D is a new random variable with the difference between X(13mg) and Y(no caffeine)

The hypothesis to be tested is H0: $\mu_{13mg} - \mu_{no\_caffeine} \leq 0$ vs H1: $\mu_{13mg} - \mu_{no\_caffeine} > 0$. For this test, the null hypothesis was rejected for an alpha of .05, which means that there is statistical evidence that the VO2max increases when athletes ingest 13mg of caffeine compared with 0mg (t = 3.2525, df = 8, p-value = .012). A 95% confidence interval for the mean difference was also obtained: [3.407, 20.01].