THE HEBREW
UNIVERSITY
OF JERUSALEM

Read the instructions **carefully** (that's a good idea in general).

- There is only a theoretical part to this one. It's a short homework – you get two weeks, but it really shouldn't take that long.

- Each person submits their own theoretical part. The theoretical part should be a single file in pdf format only (no docx or jpg) named **ex4t_ID.pdf** (ID is your ID).

- If you are submitting handwritten answers, make sure they are crystal clear.

- Points may be reduced for submissions that fail to comply.

- Make sure you follow the News forum and HW forum for any updates.

Problem 1 (Kappa).

Suppose we want to predict whether a blog post contains an argument that is liberal, conservative, or neutral. Two annotators label 900 blog posts independently and produce the following contingency table:

		$\mathcal{A}$			
		liberal	conservative	neutral	
	liberal	100	30	170	300
$\mathcal{B}$	conservative	50	130	20	200
	neutral	100	40	260	400
		250	200	450	900

(a) Compute the inter-annotator agreement in terms of percent agreement

(b) Compute Cohen's Kappa agreement under the assumption that the annotators are unbiased and therefore each one independently labels an equal number of blog posts as liberal, conservative, and neutral (i.e., 300 each).

(c) Compute the Cohen's Kappa agreement under the assumption that the annotators are biased and follow their individual biases, which are reflected in the contingency matrix.

Problem 2 (p-values).

Pick the most appropriate answer. Justify your answer.

(a) Researchers found that women who took drug X had a significant increase in bone density ($p < .05$). But women who took drug Y did not have a significant increase in bone density ($p > .05$). Which of the following statements follows?

   (1) We cannot conclude that drug X is superior to drug Y for improving bone density.

   (2) Drug X is superior to drug Y at increasing bone density.

   (3) Drug Y does not work, drug X may work.

(b) Researchers compared the nutritional habits of athletes with stress fractures to athletes without stress fractures. They considered 75 nutritional factors measured with a food frequency questionnaire. They reported that vitamin K was significantly lower ($p = .02$) in fractured athletes; and dietary fat was significantly higher ($p = .03$) in fractured athletes than non-fractured athletes. What conclusions should we draw?

(1) We should immediately begin supplementing athletes with vitamin K and recommending that they consume lower fat diets.

(2) Low vitamin K and high dietary fat intake are related to fractures in athletes, but the association may not be causal.

(3) Low vitamin K and high dietary fat intake are related to fractures in athletes, and the association is causal.

(4) The significant associations observed between vitamin K and high dietary fat and fractures are most likely chance findings.

Problem 3 (Hypothesis Testing).

A company that teaching speed reading courses claims that their students show a 5 times (500%) increase in the number of words they can read in a minute without losing comprehension. A random sample of 100 students yielded an average increase of 425%. Assume a standard deviation of 220%. Is there evidence that the company's claim is false?

(a) Perform a hypothesis test evaluating if the company's claim is reasonable or if the true average improvement is less than 500%. Use $\alpha = 0.05$ (sigh). (It's ok to use normal-distribution calculators)

(b) Calculate a 95% confidence interval for the average increase in the number of words students can read in a minute without losing comprehension.

(c) Do your results from the hypothesis test and the confidence interval agree? Explain.