Problem 6: Differing Distributions

Assigned: 3 November
Due: 18 November
Maximum Mark: 15 Points
Maximum Submission Length: 4 pages

1. Using a simulation, confirm that R's implementation of the Shapiro-Wilk normality test has a false positive rate of $\alpha=0.05$ when testing with a p-value threshold of 0.05. (i.e.: confirm that when repeatedly running a Shapiro-Wilk test on data randomly generated from a normal distribution, the null hypothesis of normality is rejected ~5% of the time.)

2. Using a simulation, estimate the *statistical power* ($1-\beta$) of the Shapiro-Wilk normality test (i.e., the proportion of the time it correctly rejects an incorrect null hypothesis) for the following scenarios, when testing at $\alpha=0.05$:

(a) A uniform distribution with sample size n=10.
(b) A uniform distribution with n=50.
(c) A uniform distribution with n=200.
(d) A Student t-distribution with 3 degrees of freedom and sample size n=10.
(e) A Student t-distribution with 3 d.o.f. and n=50.
(f) A Student t-distribution with 3 d.o.f. and n=200.

Summarize your results in a table with the following structure:

Distribution	n	$1-\beta_{SW}$
uniform	10	
uniform	50	
uniform	200	
t (dof=3)	10	
t (dof=3)	50	
t (dof=3)	200	

3. Using a simulation, estimate the statistical power ($1-\beta$) of the Komogorov-Smirnov two-sample test for distinguishing data sampled from a uniform distribution (min=-1.75,max=1.75; sample size n_1) from data sampled from a standard normal distribution ($\sigma=1$, $\mu=0$; sample size n_2), testing at $\alpha=0.05$. Calculate this for various sample sizes: $n_1 = n_2 = 20$, $n_1 = n_2 = 50$, $n_1 = n_2 = 200$, $n_1 = n_2 = 500$. Collect the results in a table as below:

n_1	n_2	$1-\beta_{KS}$
20	20	
50	50	
200	200	
500	500	

4. Repeat Question 3 above, but this time use the Anderson-Darling test instead of the K-S test. (Note: the Anderson-Darling test can be very slow, so avoid using too many trials – a precise measurement of $1-\beta$ is not crucial for this problem.) Produce a new table.

5. Based on these results, which two-sample test would you recommend, and why? Justify your conclusion quantitatively.