MATH242 HW4

Zhangir Azerbayev

Spring 2022

Problem 1

(a) For $n = 1$ we have that $\sqrt{\frac{1}{1}U_1} = \sqrt{Y^2} = |Y|$. Therefore $T = X/|Y|$. The random variables $X/|Y|$ and X/Y have the same distribution because for any $c \in \mathbb{R}$, the set of (X, Y) pairs where X/Y is c is the same as the set (X, Y) pairs where $X/|Y|$ is c.

(b) We have that

$$\begin{aligned} f(t) &= \int_{-\infty}^{\infty} |t| p_X(xt) p_Y(x) dx. \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |x| e^{-\frac{1+t^2}{2}x^2} dx \\ &= \frac{1}{\pi} \int_0^{\infty} x e^{-\frac{1+t^2}{2}x^2} dx. \\ &= \frac{1}{\pi} \frac{1}{t^2+1}. \end{aligned}$$

Because

$$\mathbb{E}(T) = \int_{\mathbb{R}} \frac{t}{t^2+1} dt$$

we have $\mathbb{E}(T) = \infty$, since the integrand is asymptotically equal to $1/t$. Similarly, since we have

$$\mathbb{E}(T^2) = \int_{\mathbb{R}} \frac{t^2}{t^2+1} dt$$

we have $\mathbb{E}(T^2) = \infty$, since the integrand is asymptotically equivalent to 1.

Problem 2

(a) Let $U_n = \sum^n Z_i^2$ where the Z_i are independent standard normals. Then $\sqrt{\frac{1}{n}U_n} = \sqrt{\frac{1}{n}\sum^n Z_i^2}$. By the law of large numbers $\frac{1}{n}\sum^n Z_i^2 \to E(Z^2) = 1$, and by the continuous mapping theorem $\sqrt{\frac{1}{n}\sum^n Z_i^2} \to \sqrt{1} = 1$.

(b) Let $Z_n = Z$ for a standard normal Z and $X_n = \sqrt{\frac{1}{n}U_n}$. Thus $Z_n \to \mathcal{N}(0,1)$ in distribution and $X_n \to 1$ in probability. Therefore by Slutsky's lemma we have $T = Z_n/X_n \to Z/1 = Z$ in distribution, therefore T approaches $\mathcal{N}(0,1)$ in distribution.

Problem 3

(a) Let $W = \sum_{i=1}^{n} kI_k$ a suggested in part (b). Due to symmetry, $I_k \sim \text{Bernoulli}(1/2)$, therefore the distribution of W does not depend on the distribution of X_i and W is pivotal.

(b) An explicit calculation shows that $\mathbb{E}(I_k) = 1/2$ and $\text{Var}(I_k) = 1/4$. Thus

$$\begin{aligned}
\mathbb{E}(W) &= \sum_{k=1}^{k} kI_k \\
&= \sum_{k=1}^{k} k\mathbb{E}(I_i) \\
&= \frac{1}{2}\sum_{k=1}^{n} k \\
&= \frac{n(n+1)}{4}.
\end{aligned}$$

Similarly

$$\begin{aligned}
\text{Var}(W) &= \sum_{k=1}^{n} k^2 \text{Var}(I_k) \\
&= \frac{1}{4}\sum_{k=1}^{n} k^2 \\
&= \frac{n(n+1)(2n+1)}{24}.
\end{aligned}$$

(c) Compute the test statistic W for you sample. Then reject the null hypothesis if $W > \Phi(1-\alpha)$.

problem4

February 23, 2022

1 Problem 4

[1]:
```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm, t, ttest_1samp, wilcoxon
```

[2]:
```
samples = np.random.standard_normal((10000, 100))
```

[3]:
```
# Likelihood ratio test

means = np.mean(samples, axis=1)

boundary = norm.ppf(.95) / np.sqrt(100)

reject_ratio = np.sum(means > boundary)/10000

print(f"Likelihood ratio test rejected {reject_ratio} of times")
```

```
Likelihood ratio test rejected 0.0524 of times
```

[4]:
```
# t-test (I can't for the life of me figure out what's wrong here)
result = ttest_1samp(samples, 0, axis=1)

tstat = result.statistic

boundary = t.ppf(.95, 99)

reject_ratio = np.sum(tstat>boundary)/10000
print(f"t test rejected {reject_ratio} of times")
```

```
t test rejected 0.0524 of times
```

[5]:
```
# Wilcoxon signed rank test
ws, ps = wilcoxon(samples, axis=1, alternative='greater')

reject_ratio = np.sum(ps<.05)/10000
print(f"wilcoxon test rejected {reject_ratio} of times")
```

```
wilcoxon test rejected 0.0544 of times
```

[6]:
```python
# Signed rank test
Ss = np.sum(samples>0, axis=1)

mean = 100 * .5
std = np.sqrt(100 * .5 * .5)

zs = (Ss-mean)/std
reject_ratio = np.sum(zs > norm.ppf(.95))/10000

print(f"sign test rejected {reject_ratio} of times")
```

```
sign test rejected 0.047 of times
```

All tests have empirical significance of around .05

1.1 Powers of tests

[7]:
```python
def get_lr_power(mean):
    samples = np.random.standard_normal((10000, 100)) + mean

    means = np.mean(samples, axis=1)

    boundary = norm.ppf(.95) / np.sqrt(100)

    reject_ratio = np.sum(means > boundary)/10000

    return reject_ratio
```

[8]:
```python
def get_t_power(mean):
    samples = np.random.standard_normal((10000, 100)) + mean

    result = ttest_1samp(samples, 0, axis=1)

    tstat = result.statistic

    boundary = t.ppf(.95, 99)

    reject_ratio = np.sum(tstat>boundary)/10000

    return reject_ratio
```

[9]:
```python
def get_wilcoxon_power(mean):
    samples = np.random.standard_normal((10000, 100)) + mean

    ws, ps = wilcoxon(samples, axis=1, alternative='greater')

    reject_ratio = np.sum(ps<.05)/10000
```

```
    return reject_ratio
```

[10]:
```
def get_sign_power(mean):
    samples = np.random.standard_normal((10000, 100)) + mean

    Ss = np.sum(samples>0, axis=1)

    mean = 100 * .5
    std = np.sqrt(100 * .5 * .5)

    zs = (Ss-mean)/std
    reject_ratio = np.sum(zs > norm.ppf(.95))/10000

    return reject_ratio
```

1.2 (B) The chart

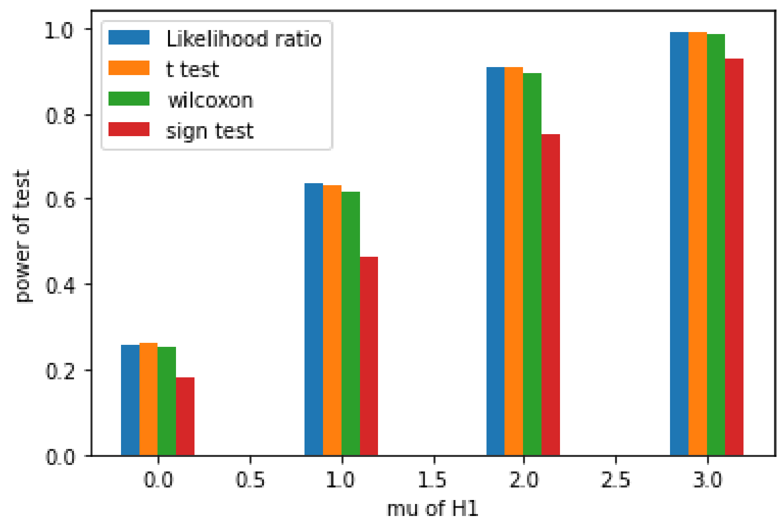
[11]:
```
mus = [0.1, 0.2, 0.3, 0.4]
lr = [get_lr_power(mu) for mu in mus]
t = [get_t_power(mu) for mu in mus]
wilc = [get_wilcoxon_power(mu) for mu in mus]
sign = [get_sign_power(mu) for mu in mus]
```

[24]:
```
x_axis = np.arange(len(mus))

plt.bar(x_axis-.15, lr, .1, label="Likelihood ratio")
plt.bar(x_axis-.05, t, .1, label="t test")
plt.bar(x_axis+.05, wilc, .1, label="wilcoxon")
plt.bar(x_axis+.15, sign, .1, label="sign test")

plt.xlabel("mu of H1")
plt.ylabel("power of test")

plt.legend()
plt.show()
```



1.3 (C)

The powers of the likelihood ratio test, t test, and wilcoxon test are close. The sign test has a much lower power than the other three. Your friend is wrong, since a test that makes greater distributional assumptions can have the advantage of being higher-powered. Our findings support Rice's conclusion. The signed rank test is almost as powerful as the t-test.