MAU11S02 Eighth (last) Monday quiz ANSWERS
Monday 12/4/21 due 4pm Monday 19/4/21

Rules and procedures. Same as always. Attempt 3 out of 6 questions. ***Remember, you must show all work.***

Question 1. X and Y are independent random variables following the distribution $B(3, 2/3)$. Last quiz, Question 4, the distributions of (X, Y) and $X + Y$ were asked. This time, calculate the mean and variance of the distribution of $X + Y$. You may give the answer directly from the definitions, or you may use properties of $E(\ldots)$ affecting the mean and variance.

Answer. The easy way is: $E(X + Y) = 2E(X) = 4$ and $\mathrm{Var}(X + Y) = 2\mathrm{Var}(X) = 2/3$.

Question 2. Count the number of numbers below, and calculate their *sample* mean or average $\overline{X}$, their *sample* variance and their *sample* standard deviation.

Note the formula is $S^2 = \frac{1}{n-1}\sum_j (X_j - \overline{X})^2$. The denominator is $n - 1$, not n. Be careful about this, even using a calculator. On a Sharp calculator, σx is the *wrong* key for sample standard deviation. The correct key is sx.

```
2.90 3.27 1.72 1.88 3.38 2.07 3.63 2.67 2.55
```

Answer.

```
n 9 Sample mean 2.6744 sample variance 0.4674 sdev 0.6837
```

Question 3. Assume that the random variables X_i sampled in question (2) are iid $N(\mu, 1)$ where μ is to be determined from the data.

Use the Normal distribution tables to estimate a 95% symmetric confidence interval for μ.

Answer.

$$\overline{X} \sim N(\mu, 1/9)$$
$$\frac{\overline{X} - \mu}{1/3} \sim N(0, 1)$$
$$\mu \in \overline{X} \mp 1.96/3$$
$$\mu \in 2.6744 \mp .6533$$
$$\mu \in [2.021, 2.328]$$

with 95% confidence.

Question 4. It is known that each random variable in the above sample is $N(\mu, \sigma^2)$ where μ and σ are to be determined from the data. That is, σ is unknown.

Use the *correct* part of the *correct* tables to calculate a 90% symmetric confidence interval for μ.

Answer. The quantity to inspect is

$$\frac{\overline{X} - \mu}{S/3} = \frac{2.6744 - \mu}{0.6837\sqrt{0.1}} = \frac{2.6744 - \mu}{0.2162}$$

Now $n = 9$, so this quantity is Student's t-distribution with 8 degrees of freedom. We want a 90% symmetric confidence interval, so we refer to the tables, percentage points at 8 degrees of freedom: the 5% point is 1.860, so the confidence interval is given: $[-1.860, 1.860]$.

The modified expression for μ belongs to this interval with 90% confidence, so the confidence interval for μ is

$$2.6744 \mp 0.2162 \times 1.860 = [1.2723, 3.0765]$$

Question 5. (These figures are fictional.) It has been believed that the average person aged between 15 and 35 makes 12 shopping trips each week.

It is *now* believed that under present conditions the average person makes more than 12 shopping trips. That the average person makes *fewer* than 12 is not considered.

From the age group, a random 100 people were consulted and the sample average (response) was 16 (trips per week) and the sample standard deviation, 2.

State the Null hypothesis and the alternative hypothesis (this is a 1-tailed test). For samples of this size, the Normal distribution may be assumed. Determine whether the outcome is significant at the 5% level.

Answer. The null hypothesis is that the average person makes 12 trips per week. The alternative hypothesis is that this is an underestimate.

$$\frac{16 - 12}{2} = 2$$
$$\text{Prob}(X \geq 16) = .0228$$

Significant at the 5% level.

Question 6. (Like question 5). A certain book was published, and the publishers believe that 70% of people who actually read the book will like it. 100 people who had read the book were questioned whether they liked or disliked the book. 63 reported that they liked it.

(i) Let X be a random variable $\sim B(100, 0.7)$. Use the *normal approximation* to estimate the probability that $X \leq 63$. (Remember the continuity correction.)

The hypothesis to test, the 'alternative hypothesis,' is that the publishers' belief — 70% like the book — is an over-estimate.

Using a 1-tailed test determine if this result is significant at (i) the 5% level, (ii) the 10% level.

Answer. The Normal approximation to the probability (with continuity correction) is the probability that

$$Y < 63.5$$

where $Y \sim N(70, 21)$, so we evaluate the probability that

$$\frac{Y - 70}{\sqrt{21}} < \frac{-6.5}{\sqrt{21}} = -1.418$$

which has probability 0.0778. This is significant at the 10% level but not at 5% level.