

Statistical Inference

Olga Klopp

QUIZ

April 2, 2021

13:30 - 15:00

Calculators are authorized.

Please turn off and store yours cell phones. Write your name and the date on the worksheet. You can use yours laptops only for the zoom session and to have access to the quiz.

- Let Φ denote the cumulative distribution function of the standard normal distribution, we have that $\Phi(2.85) = 0.998$ and $\Phi(0.285) = 0.6125$
- **Chebyshev inequality:** Let X be a random variable with finite expected value μ and finite non-zero variance σ^2 . Then for any real number $\varepsilon > 0$,

$$\mathbb{P}[|X - \mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2}$$

1. (1,5Pt) What is a statistical model? Describe a real-life application and the corresponding statistical model.
2. (1Pt) Consider the following measures: mean, median, variance, standard deviation, quantile.
 - (a) Which measures utilize relative position of the data values?
 - (b) Which measures utilize actual data values regardless of relative position?
3. (2 Pts) The *East Coast Independent News* periodically runs ads in its own classified section offering a month's free subscription to those who respond. In this way, management can get a sense about the number of subscribers who read the classified section each day. Over a period of 2 years, careful records have been kept. We assume that the number of responses per ad is a random variable which follows a distribution with mean equal to 525 and standard deviation 30. Using Chebyshev's inequality determine the interval about the mean in which at least 88.9% of the data fall.

4. Impulse Buying.

Let X represent the dollar amount spent on supermarket impulse buying in a 10-minute (unplanned) shopping interval. Based on a Denver Post article, the mean of the X distribution is about 20 and the estimated standard deviation is about 7.

- (a) (1,5 pts) Consider a random sample of $n = 100$ customers, each of whom has 10 minutes of unplanned shopping time in a supermarket. What can you say about the probability distribution of \bar{X}_n , the average amount spent by these customers due to impulse buying? What are the mean and standard deviation of \bar{X}_n ? Is it necessary to make any assumption about the \bar{X}_n distribution? Explain.
- (b) (2,5 pts) What is the probability that \bar{X}_n is between \$18 and \$22?
- (c) (2,5 pts) Let us assume that X has a distribution that is approximately normal. What is the probability that X is between \$18 and \$22?
- (d) (1 Pt) Interpretation: In part (b), we used \bar{X}_n , the average amount spent, computed for 100 customers. In part (c), we used X , the amount spent by only one customer. The answers to parts (b) and (c) are very different. Why would this happen?

- (e) (1 Pt) Almost all marketing strategies and sales pitches are designed for the average customer and not the individual customer. How does the central limit theorem tell us that the average customer is much more predictable than the individual customer?
- 5. (1 Pt) What is the Fundamental Theorem of statistics? How can you explain its meaning in simple terms?
- 6. (1 Pt) What are practical implications of Slutsky's theorem?