Definition 13.1. *Let $Y_1, \dots, Y_n$ be a* **random sample**. *The random sample ordered in an increasing order is called an* order statistic *and denoted by*

$$Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}.$$

Question Write $Y_{(1)}$ as a function of $Y_1, Y_2, \dots, Y_n$.

$$\boxed{Y_{(1)} = \min(Y_1, Y_2, \dots, Y_n)}$$

Question Write $Y_{(n)}$ as a function of $Y_1, Y_2, \dots, Y_n$.

$$\boxed{Y_{(n)} = \max(Y_1, Y_2, \dots, Y_n)}$$

Problem 13.2. *What is the distribution function of the random variable $Y_{(n)}$?*

$\longrightarrow$: for $y \in \mathbb{R}$: $F_{Y_{(n)}}(y) = \mathbb{P}[Y_{(n)} \le y] = \mathbb{P}[\max(Y_1, \dots, Y_n) \le y]$

$= \mathbb{P}[Y_1 \le y, Y_2 \le y, \dots, Y_n \le y]$ (independence)

$= \mathbb{P}[Y_1 \le y] \cdot \dots \cdot \mathbb{P}[Y_n \le y]$ (identically dist'd)

$= (\mathbb{P}[Y_1 \le y])^n = (F_Y(y))^n$

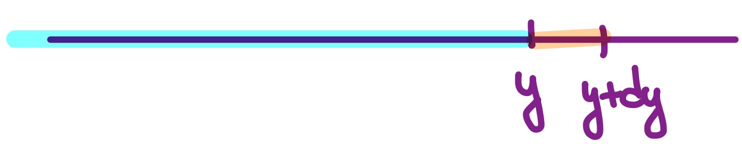
10/27/2025

Problem 13.3. *Assume that the random sample comes from a density f_Y. Is the r.v. $Y_{(n)}$ continuous? If so, what is its density $g_{(n)}$?*

$\longrightarrow$: For all y such that F_Y is differentiable:

$$g_{(n)}(y) = \frac{d}{dy} F_{Y_{(n)}}(y) = \frac{d}{dy}\left((F_Y(y))^n\right)$$

$$= n\,(F_Y(y))^{n-1} \cdot f_Y(y)$$



Problem 13.4. *What is the distribution function of the random variable $Y_{(1)}$?*

$\longrightarrow$: For $y \in \mathbb{R}$: $F_{Y_{(1)}}(y) = \mathbb{P}[Y_{(1)} \leq y] = \mathbb{P}[\min(Y_1, \dots, Y_n) \leq y]$

$= 1 - \mathbb{P}[\min(Y_1, \dots, Y_n) > y] =$

$= 1 - \mathbb{P}[Y_1 > y, Y_2 > y, \dots, Y_n > y] =$ independence

$= 1 - \mathbb{P}[Y_1 > y] \cdot \ldots \cdot \mathbb{P}[Y_n > y] =$ i.d. $= 1 - (\mathbb{P}[Y > y])^n$

$= \boxed{1 - (1 - F_Y(y))^n}$

Problem 13.5. *Assume that the random sample comes from a density f_Y. Is the r.v. $Y_{(1)}$ continuous? If so, what is its density $g_{(1)}$?*

$\longrightarrow$: For y where F_Y is differentiable:

$$g_{(1)}(y) = \frac{d}{dy} F_{Y_{(1)}}(y) = \frac{d}{dy}\left(1 - (1 - F_Y(y))^n\right)$$

$$= + n \cdot (1 - F_Y(y))^{n-1} (+1) f_Y(y)$$

$$= n \, (1 - F_Y(y))^{n-1} \cdot f_Y(y)$$

Theorem 13.2. *Let $Y_1, \dots, Y_n$ be independent, identically distributed random variables with the common cumulative distribution function F_Y and the common probability density function f_Y. Let $Y_{(k)}$ denote the k^{th} order statistic and let $g_{(k)}$ denote its probability density function. Then,*

$$g_{(k)}(y) = \frac{n!}{(k-1)!(n-k)!} (F_Y(y))^{k-1} f_Y(y) (1 - F_Y(y))^{n-k} \quad \textit{for all } y \in \mathbb{R}.$$

M378K Introduction to Mathematical Statistics
Problem Set #14
Statistics.

Definition 14.1. *A* random sample *of size n from distribution D is a random vector*

$$(Y_1, Y_2, \dots, Y_n)$$

such that

1. *$Y_1, Y_2, \dots, Y_n$ are independent, and*
2. *each Y_i has the distribution D.*

Example 14.2. Quality control. *Times until a breaker trips under a particular load are modeled as exponential. The intended procedure is to choose n breakers at random from the assembly line, subject them to the load, and measure the time it takes for them to trip. The lifetime of a specific breaker indexed by i is a random variable Y_i with an exponential distribution with an unknown parameter $\theta = \tau$. Independence of Y_i, $i = 1, \dots, n$ is assured by the random choice of breakers to test.*

Definition 14.3. *A* statistic *is a function of the (observable) random sample and known constants.*

Problem 14.1. *Give at least three examples of statistics of a certain random sample $Y_1, Y_2, \dots, Y_n$.*

- $Y_{(n)} = \max(Y_1, \dots, Y_n)$
- $Y_{(1)} = \min(Y_1, \dots, Y_n)$
- $\bar{Y} = \frac{1}{n}(Y_1 + \dots + Y_n)$ ... sample mean

Remark 14.4. *Statistics are random variables in their own right. We call their probability distributions* sampling distributions.

Example 14.5. Quality control, cont'd. *Let the random variable Y be the minimum of random variables $Y_1, \dots, Y_n$, i.e., the shortest time until the breaker is tripped in the sample. We can write*

$$Y = \min(Y_1, \dots, Y_n).$$

What is another name for this random variable? First Order Statistic

Then, the sampling distribution *of Y can be figured out by looking at its cumulative distribution function. We have …*

$$g_{(1)}(y) = n\cdot f_Y(y)\cdot (1-F_Y(y))^{n-1} = n\cdot \frac{1}{\tau} e^{-\frac{y}{\tau}} \left(\cancel{1} - (\cancel{1} - e^{-\frac{y}{\tau}})\right)^{n-1}$$

$$= \left(\frac{n}{\tau}\right) e^{-\frac{y}{\tau}} \left(e^{-\frac{y}{\tau}}\right)^{n-1} = \frac{1}{\frac{\tau}{n}}\cdot e^{-\frac{y}{\frac{\tau}{n}}}$$

$$\boxed{Y_{(1)} \sim E\left(\frac{\tau}{n}\right)}$$

Problem 14.2. *Let $Y_1, \ldots, Y_n$ be a random sample of size n from a normal distribution with mean μ and variance σ^2. What is the sampling distribution of*

$$\bar{Y}_n = \frac{1}{n}\sum_{k=1}^{n} Y_k \quad ?$$

$$\boxed{\bar{Y}_n \sim \text{Normal}\left(\text{mean} = \mu,\ \text{sd} = \frac{\sigma}{\sqrt{n}}\right)}$$

Estimators.

Def'n. The bias of an estimator $\hat{\theta}$ of the parameter θ is

$$\text{bias}(\hat{\theta}) := \mathbb{E}[\hat{\theta} - \theta]$$

Notation from the "book": "$\mathbb{E}_\theta(\cdot)$, $\mathbb{E}^\theta(\cdot)$, $\mathbb{E}[\cdots / \theta]$"

We say that an estimator $\hat{\theta}$ is unbiased for the parameter θ if

$$\boxed{\text{bias}(\hat{\theta}) = 0 \iff \mathbb{E}[\hat{\theta}] = \theta}$$

for all possible values of θ.