

## 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  $Y_1, Y_2, \dots, Y_n$ .

**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?

Then, the sampling distribution of  $\bar{Y}$  can be figured out by looking at its cumulative distribution function. We have ...

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

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