

## TASK Probability

Mosleh Abdelsattar

- 1] It's a random variable that can take only one of two values (e.g. 0 or 1)
  - Exmp: A student pass or fail a specific course
- 2] Binomial is just Bernoulli repeated multiple times
  - use Bernoulli when # trials is only one
  - use Binomial when # trials is more than one and trials are independent from each other
- 3] They're totally unrelated (independent)
- 4] Discrete: the variable can take finite # values
  - Ex: # people working in a company
  - Continuous: the variable can take infinite # values
  - Ex: temperature degree
- 5] 
$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$n=10, x=9, p=0.9$
- 6] test of independence
  - Null hypothesis: they're independent
  - Alternative hypothesis: they're not independent
- 7] Because covariance value depends on the scale of variables and it measures how variables vary together. You still need correlation.
- 8] PDF: relation between continuous random variable and its probability
  - The PDF by itself is not so useful until integrated over an area.
  - CDF: integrates from  $-\infty$  to the variable of interest
  - $$CDF = \int_{-\infty}^x f(x) dx$$
 → PDF
- 9]  $p < \alpha$ , then reject null hypothesis
  - It means "the traffic to the App is not evenly distributed across weekdays"
- 10] Assumption: successive replies are independent
  - $X = 5 + 3$       what's  $X$ ?      reply is 9 (wrong answer)  
                         multiplying  $X$  by 2      reply is 18 (also wrong answer)

- 11
- Correlation  $\neq$  Causation, there may be other factors
  - Motivation. A motivated student may get higher grades
  - we can't say from correlation alone that time on platform is the direct cause of higher grades

- 12
- I would integrate the PDF from 5 mins to  $\infty$  time
  - or  $1 - \text{CDF}(5 \text{ mins})$