① If the sample means differ for each ~~group~~ other, you cannot conclude that ~~the~~ any of the groups differ from the population. The Central Limit Theorem (CLT) backs this up. All sample means ultimately approximate to the population mean.

② Statistical Estimation: Making inferences on the population from facts derived from the sample

Hypothesis Testing: Making statistical decisions on a H_0 which states that there is no difference between the sample & the population.

③ $P(\text{one die is } 3 \mid \text{sum}=6) = \dfrac{(3,3)}{(1,5),\ (2,4)\ (3,3)\ (4,2)\ (5,1)} = \boxed{\dfrac{1}{5}}$

④ 7 black, 8 blue, 9 green.

a $P(\text{match}) = \underset{\text{black}}{\dfrac{7}{96}} + \underset{\text{blue}}{\dfrac{7}{72}} + \underset{\text{green}}{\dfrac{1}{8}} = \dfrac{85}{288} = \boxed{29.5\%}$

b $P(\text{Black}) = \boxed{\dfrac{7}{96}}$

(5) $P(\text{distinct digits}) = \frac{7!}{10^7} = \frac{846}{10^7}$

$= \boxed{8.46 \times 10^{-5}}$

(6) $P(x=1) = p$
$P(x=0) = 1-p$

$\mu = E(x) = 1 \cdot p + (1-p) = p$

$\sigma^2 = E(x-\mu)^2 = (1-p)^2 p + (0-p)^2(1-p) = p(1-p)$
$\Rightarrow \sigma = \sqrt{p(1-p)}$

(7) PMF is used to calculate the point probability of a discrete rv, whereas a PDF is used to calculate point probability of a continuous rv.

(8) CLT tells us what the true statistics of the population are, or will be. Given the guarantee that, through repeated sampling, that the statistics will take on the shape of the population statistics, allow us to identify trends & anomalies.

CLT is useful due it always approximating to normal.

(9) Deviations from the sample mean are usually less than deviations from the true population mean. In order to compensate for underestimates we use n-1 degrees of freedom for the divisor.

(10) A 100% confidence interval would be so wide that it would contain alot of useless data, as well as statistical aberrations & outliers. Continuous rv's can sometimes have infinite range, so a 100% CI would be $[-\infty, \infty]$, which is useless. We wish to narrow down our range of data w/ CI's, so a 100% CI is counterproductive to that task.

(11) P-value of one-sided test > P-val of 2-sided test. We are forced to use $\alpha/2$ for 2 sided tests b/c the values can be on either side of the curve.

(12) Failing to reject H_0 simply means there is not enough data to reject the claim made by H_0. The possibility of a different sample rejecting H_0 still exists.

(13) $f(x_n) = \lambda e^{-\lambda x_n} = f(x)$

$$L(f(x)) = \sum_{n=0}^{n-1} \log(f(x_n))$$

$$\log(f x_n) = \log \lambda e^{-\lambda x_n} = \log \lambda - \lambda x_n$$

$$\Rightarrow L(f(x)) = N \log \lambda - \lambda \sum x_n$$

$$\frac{\partial}{\partial \lambda} L(f(x)) = 0 \Rightarrow \frac{N}{\lambda} - \sum_{n=0}^{n-1} x_n = 0$$

$$\Rightarrow \hat{\lambda} = \frac{1}{\frac{\sum_{n=0}^{n-1} x_n}{n}} = \frac{1}{\bar{Y}}$$

(14) Multiple comparisons occur when a set of statistical inferences infers a subset of parameters based on the observed values.

(15) Bonferroni multiple comparisons sets significance cut-off @ α/n. Explicitly, given n tests, for hypothesis H_i $(1 \leq i \leq n)$ under assumption of H_0 that $\forall H_i$ = false, & if test criteria $\leq \alpha/n$, $\rightarrow$ crit val $= \alpha$

(16) B & D

(17) A & D

(18) H_0: $p_M = 0.25$
$p_N = 0.5$

H_1: $p_M \neq 0.25$ OR
$p_N \neq 0.5$

$\chi^2 = 6.53 \Rightarrow$ p-val $= 0.038 < \alpha = 0.05$

Reject H_0, at least one of p_M or p_N is wrong

(19) a H_0: $\beta_1 = 0$

H_1: $\beta_1 \neq 0$

b No, a strong linear relationship is marked by a small SE, which this does not have.
~~Also Bigger~~

T-statistic $= \dfrac{22.464}{53.06} = 0.428$

$\Rightarrow$ p-val $= 2 \cdot (0.3378)$
$= 0.6757 > 0.05$

DF $= 48$

$\boxed{\text{Fail to reject } H_0}$

(20) We can conclude, w/ 90% confidence, that the true population mean is between 37.5 & 40.