

統計學

共五題，總計 100 分。依序作答。

1. (25%) A bag contains 5 balls with different colors. Draw 10 balls one by one with replacement. Let $X_i = 1$ if color i appears in the 10 draws; otherwise, $X_i = 0$, for $i = 1, 2, \dots, 5$. Let N denote the number of colors observed in the 10 balls drawn.
 - (a) (5%) What is the probability that color 1 does not appear in the first draw?
 - (b) (5%) Give the expression (instead of the numerical value) for $P(X_i = 0)$.
 - (c) (5%) Express N in terms of X_i 's.
 - (d) (5%) Give the expression (instead of the numerical value) for $E(N)$.
 - (e) (5%) Give the expression (instead of the numerical value) for the average number of draws required to observe all 5 colors.

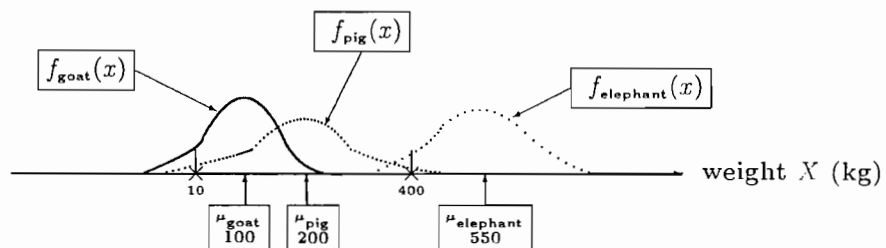
2. (15%) Suppose that random sample $\{X_1, \dots, X_{25}\}$ is drawn from $N(\mu, \sigma^2 = 2^2)$ with μ unknown. The goal is to test $H_0 : \mu = 30$ against $H_1 : \mu > 30$. If the *rejection region* is

$$\left\{ (x_1, \dots, x_{25}) \mid \bar{x} = \sum_{i=1}^{25} x_i / 25 \geq 30.8 \right\}.$$

- (a) (7%) Find the significance level of the test.
- (b) (8%) Find the *power* of this test at $\mu = 31.2$.

(You may need $P(|Z| < 1) = 68\%$, $P(|Z| < 2) = 95\%$, where Z is the standard normal random variable.)

3. (10%) The following diagram shows the probability density functions of the weights of pig, goat and elephant, where μ_k denotes the mean weight of animal k .



- (a) (5%) What kind of animal would you guess if its weight is 10 kilograms? Explain.
- (b) (5%) What kind of animal would you guess if its weight is 400 kilograms? Explain.

4. (20%) Pollution of the rivers in Taiwan has been a problem for many years. Consider the following events:

A = The river is polluted

\bar{A} = The river is not polluted

B = A sample of water tested detects pollution

C = Fishing is permitted

Suppose that $P(A) = 0.3$, $P(B|A) = 0.75$, $P(B|\bar{A}) = 0.2$, and $P(C|A \cap B) = 0.2$.

(a) (5%) Find $P(A \cap B)$.

(b) (5%) Find $P(A \cap B \cap C)$.

(c) (5%) Find $P(B)$.

(d) (5%) Find $P(A|B)$.

5. (30%) An analyst is interested in computing a pooled estimate of the variance of two populations. Pooling is justified if it appears reasonable to assume that the two populations have equal variances $\sigma_1^2 = \sigma_2^2 = \sigma_p^2$. The following data are collected.

$$n_1 = 8, S_1^2 = 40.84, n_2 = 8, S_2^2 = 86.50$$

(a) (10%) Test if $\sigma_1^2 = \sigma_2^2$? (with $\alpha = 0.1$)

(b) (10%) What is the pooled estimate of σ_p^2 ?

(c) (10%) Compute a 95% confidence interval for σ_p^2 .

(Note that $F_{0.05,7,7} = 3.79$, $F_{0.95,7,7} = 0.2639$, $\chi_{0.025,14}^2 = 26.119$, and $\chi_{0.975,14}^2 = 5.62872$.)