Statistical Inference
Welcome Tutorial :-)
Tutorial 1

Lecturer: Ming Gao

DaSE @ ECNU

May. 5, 2018

Tutorial 1

1. Prove each of the following statements:
   a. If $P(B) = 1$, then $P(A|B) = P(A)$ for any A;
   b. If $A \subset B$, then $P(B|A) = 1$ and $P(A|B) = \frac{P(A)}{P(B)}$;
   c. If A and B are mutually exclusive, then

$$P(A|A \cup B) = \frac{P(A)}{P(A) + P(B)}.$$

   d. $P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C)$.
2. A certain river floods every year. Suppose that the low-water mark is set at 1 and the high-water mark Y has distribution function

$$F_Y(y) = P(Y \leq y) = 1 - \frac{1}{y^2}, 1 \leq y \leq \infty.$$

   a. Verify that $F_Y(y)$ is a cdf;
   b. Find $f_Y(y)$, the pdf of Y;

3. Let X have the pdf

$$f(x) = \frac{1}{2}(1+x), -1 < x < 1.$$

   a. Find the pdf of $Y = X^2$;
   b. Find $E(Y)$ and $Var(Y)$;
   c. Find $E(aY+b)$ and $Var(aY+b)$, where a and b are constant.

4. In each of the following find the pdf of Y. Show that the pdf integrates to 1.
   a. $Y = X^3$ and $f_X(x) = 42x^5(1-x), 0 < x < 1$;
   b. $Y = 4X+3$ and $f_X(x) = 7e^{-7x}, 0 < x < \infty$;
   c. $Y = X^2$ and $f_X(x) = 30x^2(1-x)^2, 0 < x < 1$;
   d. $Y = X^2$ and $f_X(x) = 1, 0 < x < 1$.

5. For each of the following families, please verify that it is an exponential family:
   a. $N(\mu, \sigma^2)$;
   b. $N(\theta, a\theta)$, a known;
   c. $f(x|\theta) = C \cdot exp^{\left(-(x-\theta)^4\right)}$, C is a normalizing constant.
6. Calculate $P(|X - \mu_X| \geq k\sigma_X)$ for $X \sim uniform(0,1)$ and $X \sim exponential(\lambda)$, and compare your answers to the bound from Chebychev's inequality and Chernoff bound.
7. A pdf is defined by

$$f(x,y) = \begin{cases} C(x+2y), & \text{if } 0 < y < 1 \text{ and } 0 < x < 2 \\ 0, & \text{otherwise.} \end{cases}$$

   a. Find the value of C;
   b. Find the marginal distribution of X;
   c. Find the joint cdf of X and Y;
   d. Find the pdf of the r.v. $Z = \frac{9}{(X+1)^2}$.

8. a. Find $P(X > \sqrt{Y})$ if X and Y are jointly distributed with pdf
$$f(x,y) = x + y, 0 \leq x \leq 1, 0 \leq y \leq 1.$$
   b. Find $P(X^2 < Y < X)$ if X and Y are jointly distributed with pdf
$$f(x,y) = 2x, 0 \leq x \leq 1, 0 \leq y \leq 1.$$
9. Prove that if the joint cdf of X and Y satisfies
$$F_{X,Y}(x,y) = F_X(x)F_Y(y),$$
then for any pair of intervals (a,b) and (c,d),
$$P(a \leq X \leq b, c \leq Y \leq d) = P(a \leq X \leq b)P(c \leq Y \leq d).$$
10. Let $X \sim N(\mu, \sigma^2)$ and let $Y \sim N(\gamma, \sigma^2)$. Suppose X and Y are independent. Define $U = X + Y$ and $V = X - Y$. Show that U and V are independent normal r.v.s. Find the distribution of each of them.

11. Let X and Y be independent r.v.s with means μ_X, μ_Y and variances σ_X^2, σ_Y^2. Find an expression for the correlation of XY and Y in terms of these means and variances.
12. Let X_1, X_2, and X_3 be uncorrelated r.v.s, each with mean μ and variance σ^2. Find $Cov(X_1 + X_2, X_2 + X_3)$ and $Cov(X_1 + X_2, X_1 - X_2)$.
13. Let $X_1, \cdots, X_n$ be i.i.d. r.v.s with continuous cdf F_X, and suppose $E(X_i) = \mu$. Define the r.v.s

$$Y_i = \begin{cases} 1, & \text{if } X_i > \mu, \\ 0, & otherwise. \end{cases}$$

Find $E(Y_i), Var(Y_i)$, and the distribution of $\sum_{i=1}^{n} Y_i$.

14. Establish the following recursion relations for means and variances. Let $\overline{X}_n$ and S_n^2 be the mean and variance, respectively, of $X_1, \cdots, X_n$. Then suppose another observation X_{n+1}, becomes available. Show that
    a. $\overline{X}_{n+1} = \frac{X_{n+1} + n\overline{X}_n}{n+1}$;
    b. $nS_{n+1}^2 = (n-1)S_n^2 + \frac{n}{n+1}(X_{n+1} - \overline{X}_n)^2$.
15. Let $X_1, \cdots, X_n$ be a random sample from a population with pdf

$$f_X(x) = \begin{cases} \frac{1}{\theta}, & \text{if } 0 < x < \theta; \\ 0, & \text{otherwise.} \end{cases}$$

    Let $X_{(1)} < \cdots < X_{(n)}$ be the order statistics. Show that $\frac{X_{(1)}}{X_{(n)}}$ and $X_{(n)}$ are independent r.v.s.
16. If $\overline{X}_1$ and $\overline{X}_2$ are the mean of two independent samples of size n from a population with variance σ^2, find a value for n so that $P(|\overline{X}_1 - \overline{X}_2| < \frac{\sigma}{5}) \approx 0.99$. Justify your calculations.

17. Let $X_1, \cdots, X_n$ be a random sample from a population with mean μ and variance σ^2. Show that

$$E\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}\right) = 0 \text{ and } Var\left(\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}\right) = 1.$$

Thus, the normalization of $\overline{X}_n$ in the central limit theorem gives r.v.s that have the same mean and variances as the limiting $N(0,1)$ distribution.

18. Let $X_1, \cdots, X_n$ be a random sample from the pdf

$$f(x|\mu,\sigma) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma}, \mu < x < \infty, 0 < \sigma < \infty.$$

Find a two-dimensional sufficient statistic for (μ, σ).

19. For each of the following distributions let $X_1, \cdots, X_n$ be a random sample. Find a minimal sufficient statistic for θ
    a. $f(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}, x \in R, \theta \in R;$
    b. $f(x|\theta) = e^{-(x-\theta)}, \theta < x < \infty, \theta \in R.$

20. One observation, X, is taken from a $N(0, \sigma^2)$ population.
    a. Find an unbiased estimator of σ^2, and justify your answer;
    b. Find the MLE of σ.
21. Let $X_1, \cdots, X_n$ be a random sample from a population with pmf

$$P_\theta(X = x) = \theta^x (1-\theta)^{1-x}, x = 0 \text{ or } 1, 0 \leq \theta \leq \frac{1}{2}$$

    a. Find the method of moments estimator and MLE of θ;
    b. Find the mean squared errors of each of the estimators;
    c. Which estimator is preferred? Justify your choice.
22. Let $X_1, \cdots, X_n$ be i.i.d. *Poisson*(λ), and λ have a *gamma*(α, β) distribution, the conjugate family for the Poisson
    a. Find the posterior distribution of λ;
    b. Calculate the posterior mean and variance.

23. Let $X_1, \cdots, X_n$ be i.i.d. *Bernoulli*(p). Show that the variance of $\overline{X}$ attains the Cramér-Rao lower bound, and hence $\overline{X}$ is the best unbiased estimator of p.
24. Let $X_1, \cdots, X_n$ be a random sample from a population with mean μ and variance σ^2
    a. Show that the estimator $\sum_{i=1}^{n} a_i X_i$ is an unbiased estimator of μ if $\sum_{i=1}^{n} a_i = 1$;
    b. Among all unbiased estimators of this form find the one with minimum variance, and calculate the variance.