TEST CODE: PSA (Objective type) and PSB (Short answer type) 2016

SYLLABUS

Mathematics

Arithmetic, geometric and harmonic progressions. Trigonometry. Two dimensional coordinate geometry: Straight lines, circles, parabolas, ellipses and hyperbolas.

Elementary set theory. Functions and relations. Elementary combinatorics: Permutations and combinations, Binomial and multinomial theorem.

Theory of equations.

Complex numbers and De Moivre's theorem.

Vectors and vector spaces. Algebra of matrices. Determinant, rank, trace and inverse of a matrix. Solutions of linear equations. Eigenvalues and eigenvectors of matrices.

Limits and continuity of functions of one variable. Differentiation. Leibnitz formula. Applications of differential calculus, maxima and minima. Taylor's theorem. Indefinite integral. Fundamental theorem of calculus. Riemann integration and properties. Improper integrals.

Statistics and Probability

Notions of sample space and probability. Combinatorial probability. Conditional probability and independence. Bayes Theorem. Random variables and expectations. Moments and moment generating functions. Standard univariate discrete and continuous distributions. Distribution of functions of a random variable. Distribution of order statistics. Joint probability distributions. Marginal and conditional probability distributions. Multinomial distribution. Bivariate normal and multivariate normal distributions.

Sampling distributions of statistics. Statement and applications of Weak law of large numbers and Central limit theorem.

Descriptive statistical measures. Contingency tables and measures of association. Product moment and other types of correlation. Partial and multiple correlation. Simple and multiple linear regression.

Elementary theory of estimation (unbiasedness, minimum variance, sufficiency). Methods of estimation (maximum likelihood method, method of moments). Tests of hypotheses (basic concepts and simple applications of Neyman-Pearson Lemma). Confidence intervals. Inference related to regression. ANOVA. Elements of nonparametric inference.

Basic experimental designs such as CRD, RBD, LSD and their analyses. Elements of factorial designs. Conventional sampling techniques (SRSWR/SRSWOR) including stratification. Ratio and regression methods of estimation.

SAMPLE QUESTIONS: PSA

Note: For each question there are four suggested answers of which only one is correct.

1. The number of functions $f : \{1, 2, \ldots, 10\} \to \{1, 2, \ldots, 10\}$ such that $f(x) \neq x$ for all x is

   (A) 10! (B) 9^{10} (C) 10^9 (D) $10^{10} - 1$.

2. The set of all ordered pairs of real numbers (x, y) satisfying satisfying $y^2 - 2y - x^2 + 4x = 3$ is a

   (A) circle (B) point (C) hyperbola (D) pair of straight lines.

3. Let
$$f(x) = \lim_{n\to\infty} \frac{\log(2 + x) - x^{2n} \sin x}{1 + x^{2n}} \text{ for } x > 0.$$
   Then

   (A) f is continuous at $x = 1$
   (B) $\lim_{x\to 1+} f(x) \neq \lim_{x\to 1-} f(x)$
   (C) $\lim_{x\to 1+} f(x) = \sin 1$
   (D) $\lim_{x\to 1-} f(x)$ does not exist.

4. Suppose a real matrix A satisfies $A^3 = A$, $A \neq I$, $A \neq 0$. If $\text{Rank}(A) = r$ and $\text{Trace}(A) = t$, then

   (A) $r \geq t$ and $r + t$ is odd
   (B) $r \geq t$ and $r + t$ is even
   (C) $r < t$ and $r + t$ is odd
   (D) $r < t$ and $r + t$ is even.

5. Let
$$f(x) = x^2 + \frac{1}{x^2} + x + \frac{1}{x}, \quad x > 0$$
   and let $m = \min\{f(x)\}$. Then

   (A) $m = 1$ (B) $m = 4$ (C) $m = 27/4$ (D) m does not exist.

6. Let f be a convex function, i.e.,

$$f(tx+(1-t)y) \leq tf(x)+(1-t)f(y)$$

for all $0 \leq t \leq 1$ and $x, y \in \mathbb{R}$. Then which of the following is necessarily true?

(A) $2f(0)+f(4) \geq 2f(1)+f(2)$

(B) fg is a convex function whenever g is convex

(C) f is nondecreasing

(D) none of these.

7. Suppose A is a 100×100 real symmetric matrix whose diagonal entries are all positive. Then which of the following is necessarily true?

(A) All eigenvalues of A are greater than 0

(B) no eigenvalue of A is greater than 0

(C) at least one eigenvalue of A is greater than 0

(D) none of these.

8. The integral

$$\int_0^1 \frac{\sin x}{x^\alpha}\, dx$$

(A) is finite only for $\alpha = 0$

(B) is finite only for $|\alpha| < 1$

(C) is finite for all $\alpha < 2$

(D) is infinite for any value of α.

9. Given θ in the range $0 \leq \theta < \pi$, the equation

$$2x^2+2y^2+4x\cos\theta+8y\sin\theta+5=0$$

represents a circle for all θ in the interval

(A) $0 < \theta < \pi/3$ (B) $\pi/4 < \theta < 3\pi/4$

(C) $0 < \theta < \pi/2$ (D) $0 \leq \theta < \pi$.

10. How many 5×5 matrices are there such that each entry is 0 or 1 and each row sum and each column sum is 4?

(A) 64 (B) 32 (C) 120 (D) 96.

11. For $n \geq 1$, let G_n be the geometric mean of $\{\sin(\frac{\pi}{2} \cdot \frac{k}{n}) : 1 \leq k \leq n\}$. Then $\lim\limits_{n \to \infty} G_n$ is

(A) $1/4$ (B) $\log 2$ (C) $\frac{1}{2} \log 2$ (D) $1/2$.

12. Suppose a, b, x, y are real numbers such that $a^2 + b^2 = 81$, $x^2 + y^2 = 121$ and $ax + by = 99$. Then the set of all possible values of $ay - bx$ is

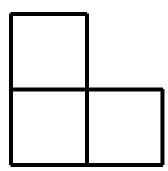
(A) $\{0\}$ (B) $\left(0, \frac{9}{11}\right]$ (C) $\left(0, \frac{9}{11}\right)$ (D) $\left[\frac{9}{11}, \infty\right)$.

13. In a triangle with sides of length a, b, c, suppose $b + c = x$ and $bc = y$. If also $(x + a)(x - a) = y$, then the triangle is necessarily

(A) equilateral (B) right angled

(C) acute angled (D) obtuse.

14. Three distinct squares are selected at random from a 8×8 chess board. Then the probability that they form an L-shaped pattern (looked at from one fixed side only) as drawn below is



(A) $\frac{196}{\binom{64}{3}}$ (B) $\frac{49}{\binom{64}{3}}$ (C) $\frac{36}{\binom{64}{3}}$ (D) greater than $1/2$.

15. Suppose X is distributed as Poisson with mean λ. Then $E(1/(X + 1))$ is

(A) $\frac{e^{\lambda} - 1}{\lambda}$ (B) $\frac{1}{\lambda + 1}$ (C) $\frac{1 - e^{-\lambda}}{\lambda}$ (D) $\frac{1 - e^{-\lambda}}{\lambda + 1}$.

16. A permutation of $1, 2, \ldots, 100$ is chosen at random. Then the probability that the numbers 1 and 100 appear next to each other equals

(A) $1/100$ (B) $1/50$ (C) $1/99$ (D) $1/98$.

17. There are 10 boxes each containing 6 white and 7 red balls. Two different boxes are chosen at random, one ball is drawn simultaneously at random from each and transferred to the other box. Now a box is again chosen from the 10 boxes and a ball is chosen from it. Then the probability that this ball is white is

(A) 6/13 (B) 7/13 (C) 5/13 (D) none of these.

18. The number of cars (X) arriving at a service station per day is a random variable with mean 4. The service station can provide service to a maximum of 4 cars per day. Then the expected number of cars per day that do not get serviced equals

(A) 4 (B) 0 (C) $\sum_{i=0}^{\infty} iP(X = i + 4)$ (D) $\sum_{i=4}^{\infty} iP(X = i - 4)$.

19. Suppose X_1 and X_2 are independent random variables distributed as $\mathrm{Ber}(p_1)$ and $\mathrm{Ber}(p_2)$ respectively. Then $Y = \max(X_1, X_2)$ is distributed as a Bernoulli random variable with success probability

(A) $1 - p_1p_2$ (B) $p_1 + p_2 - p_1p_2$ (C) $\max\{p_1, p_2\}$ (D) $\min\{1 - p_1, 1 - p_2\}$.

20. Assume that (X, Y) is bivariate Normal with $E(X) = E(Y) = 0$, $\mathrm{Var}(X) = \sigma_1^2$, $\mathrm{Var}(Y) = \sigma_2^2$ and $\mathrm{Cor}(X, Y) = \rho$ for some $\rho \in (-1, 1)$. The probability that X is larger than Y is

(A) 1/2 (B) $\Phi\left(\frac{\sigma_1-\sigma_2}{\sqrt{1-\rho^2}}\right)$ (C) $\Phi\left(\frac{\sigma_1^2+\sigma_2^2}{\sqrt{1-\rho^2}}\right)$ (D) none of the above.

21. Suppose that the bivariate data $(x_1, y_1), \ldots, (x_n, y_n)$ lie on the straight line $y = a + bx$ for some $a, b \in \mathbb{R}$. Assume further that neither all the x_i's are same, nor are all the y_i's. Which of the following values is not a possibility for the correlation coefficient calculated from the above data?

(A) 1 (B) -1 (C) 0 (D) None of the above.

22. In the randomised block design for ANOVA where k is the number of treatments and b is the number of blocks, the degrees of freedom for error is given by

(A) $bk - 1$ (B) $kb + 1$ (C) $(b - 1)(k - 1)$ (D) $k + b - 1$.

23. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a population with mean μ and finite variance. Consider the following two estimators of μ^2:

$$\begin{aligned} E_1 &= \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \sum X_i X_j \\ E_2 &= \bar{X}^2 - \frac{1}{n(n-1)} \sum_{1 \leq i \leq n} (X_i - \bar{X})^2. \end{aligned}$$

Then,

(A) E_1 is unbiased but E_2 is biased

(B) E_1 is biased but E_2 is unbiased

(C) both E_1 and E_2 are unbiased

(D) both E_1 and E_2 are biased.

24. It is known that the proportion of smokers (p) in a population lies in the interval $[1/3, 2/3]$. In a random sample of N individuals selected from the population, it was found that M were smokers. The maximum likelihood estimate of p based on the above data is

(A) $\max\{1/3, M/N\}$

(B) $\min\{M/N, 2/3\}$

(C) M/N

(D) none of the above.

25. Suppose that the least squares linear regression equation of y on x is $y = a + bx$ and that of x on y is $x = c + dy$. If it is known that $b \neq 0$ and $d \neq 0$, then the ratio of the standard deviation of x to the standard deviation of y

(A) is $\frac{1}{\sqrt{bd}}$

(B) is $\sqrt{d/b}$

(C) is $\sqrt{b/d}$

(D) cannot be determined from the above information.

26. Let $\lfloor x \rfloor$ denote the largest integer not larger than x. If X is distributed as $N(0,1)$, then $E(\lfloor X \rfloor)$

(A) is 0 (B) is 0.5 (C) is -0.5 (D) does not exist.

27. X and Y are i.i.d. random variables with finite variances. Then

(A) $Var(XY) = Var(X)Var(Y)$

(B) $Var(XY) \geq Var(X)Var(Y)$

(C) $Var(XY) \leq Var(X)Var(Y)$

(D) none of the above is necessarily true.

28. Suppose $X_1, X_2, \ldots, X_n$ is a random sample of size n from the probability density

$$f(x) = \frac{\alpha^p}{\Gamma(p)} x^{p-1} e^{-\alpha x}\,, x > 0$$

where p is a known positive constant and $\alpha > 0$ is an unknown parameter. Let $\hat{\alpha} = p/\bar{X}$ be a proposed estimator of α. Then,

(A) $E(\hat{\alpha}) = \alpha$

(B) $E(\hat{\alpha}) = \frac{\alpha}{1-\frac{1}{np}}$

(C) $E(\hat{\alpha}) = \frac{\alpha}{1-np}$

(D) none of the above statements is true.

29. The sign test is a nonparametric procedure for testing

(A) whether two populations have identical means.

(B) whether two populations have identical medians.

(C) whether two populations have identical probability distributions.

(D) whether two populations are independent.

30. Let $X_1, \ldots, X_n$ be i.i.d. from $N(0, \sigma^2)$. What is the form of the most powerful test for testing the null hypothesis

$$H_0 : \sigma = 1\,,$$

against the alternative

$$H_1 : \sigma = 2\,?$$

(A) Reject H_0 if $\sum_{i=1}^n X_i^2 > c$

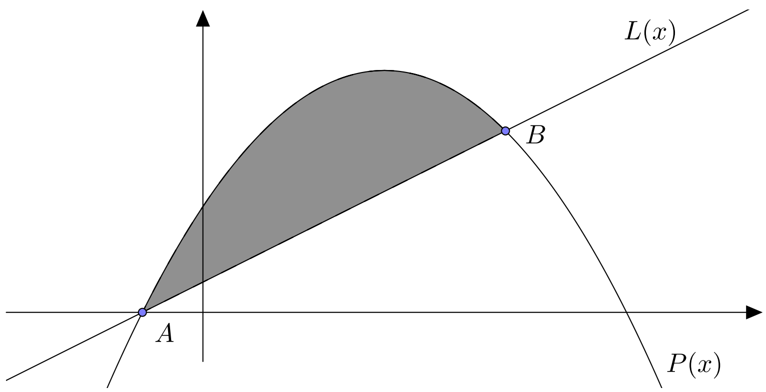
(B) Reject H_0 if $\sum_{i=1}^n X_i^2 < c$

(C) Reject H_0 if $\sum_{i=1}^n (X_i - \bar{X})^2 > c$

(D) Reject H_0 if $\sum_{i=1}^n (X_i - \bar{X})^2 < c$.

SAMPLE QUESTIONS: PSB

1. In the diagram below, $L(x)$ is a straight line that intersects the graph of a polynomial $P(x)$ of degree 2 at the points $A = (-1, 0)$ and $B = (5, 12)$. The area of the shaded region is 36 square units.
   Obtain the expression for $P(x)$.



2. Let $f : [-1, 1] \to \mathbb{R}$ be a continuous function. Suppose that $f'(x)$ exists and $f'(x) \leq 1$ for all $x \in (-1, 1)$. If $f(1) = 1$ and $f(-1) = -1$, prove that

$$f(x) = x \text{ for all } x \in [-1, 1].$$

3. Suppose A is an $n \times n$ real symmetric matrix such that

$$Tr(A^2) = Tr(A) = n.$$

   Show that all the eigenvalues of A are equal to 1.

4. For each $c \in \mathbb{R}$, define a function $T_c \colon \mathbb{R}^4 \to \mathbb{R}^4$ by

$$T_c\big(x_1, x_2, x_3, x_4\big) := \big((1+c)x_1, x_2 + cx_3, x_3 + cx_2, (1+c)x_4\big).$$

   For every $c \in \mathbb{R}$, find the dimension of the null space of T_c.

5. A box contains 50 red balls, 30 green balls and 20 blue balls. Suppose balls are drawn successively at random with replacement from the box. Let N denote the minimum number of draws required to obtain balls of all three colours. Compute $P(N > n)$ for all positive integers n.

6. Let $X_1, \ldots, X_n$ be i.i.d. random variables from a continuous distribution whose density is symmetric around 0. Suppose $E(|X_1|) = 2$. Define

$$Y = \sum_{i=1}^{n} X_i \quad \text{and} \quad Z = \sum_{i=1}^{n} \mathbf{1}(X_i > 0)\,.$$

Calculate the covariance between Y and Z.

7. Suppose $\{(y_i, x_{1i}, x_{2i}, \ldots, x_{ki}) : i = 1, 2, \ldots, n_1 + n_2\}$ represents a set of multivariate observations. It is found that the least squares linear regression fit of y on $(x_1, \ldots, x_k)$ based on the first n_1 observations is the **same** as that based on the remaining n_2 observations, and is given by

$$y = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_j\,.$$

If the regression is now performed using all $(n_1 + n_2)$ observations, will the regression equation remain the same? Justify your answer.

8. Suppose that $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ are the coordinates of n points chosen independently and uniformly at random within a circle with centre $(0, 0)$ and unknown radius r.
(a) Obtain the MLE $\hat{r}_n$ of r.
(b) Examine whether $\hat{r}_n$ is sufficient for r.
(c) For any $\varepsilon > 0$, show that $\lim_{n\to\infty} P(|\hat{r}_n - r| > \varepsilon) = 0$.

9. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from an exponential distribution with mean λ. Assume that the observed data is available on $[X_1], \ldots, [X_n]$, instead of $X_1, \ldots, X_n$, where $[x]$ denotes the largest integer less than or equal to x. Consider a test for $H_0\colon \lambda = 1$ vs $H_1\colon \lambda > 1$ which rejects H_0 when $\sum_{i=1}^{n} [X_i] > c_n$. Given $\alpha \in (0, 1)$, obtain values of c_n such that the size of the test converges to α as $n \to \infty$.

10. A cake weighing one kilogram is cut into two pieces, and each piece is weighed separately. Denote the measured weights of the two pieces by X and Y. Assume that the errors in obtaining X and Y are independent and normally distributed with mean zero and the same (unknown) variance. Devise a test for the hypothesis that the true weights of the two pieces are equal.