Math 3339 – Fall 2017 Name:

Final Exam

Please, write clearly and justify all your steps, to get proper credit for your work. You are allowed to use the textbook and your hand calculator. You are supposed to use the statistical tables from the book to solve some of the problems.

(1) [4.5 Pts] From an ordinary deck of 52 playing cards, cards are drawn successively at random and without replacement. Compute:

(a) The probability that the first five cards drawn from the deck contain at most 1 spade.

(b) The probability that the first five cards drawn from the deck contain exactly 2 spades.

(c) The probability that the third spade appears on the sixth draw.

SOLUTION

(a) $P(\text{at most 1 spade}) = P(\text{no spades}) + P(\text{exactly 1 spade}) = \frac{\binom{39}{5}}{\binom{52}{5}} + \frac{\binom{13}{1}\binom{39}{4}}{\binom{52}{5}}$

(b) $P(\text{exactly 2 spade}) = \frac{\binom{13}{2}\binom{39}{3}}{\binom{52}{5}}$

(c) $P(2\spadesuit \text{ in 5 draws})P(\spadesuit \text{ on sixth draw}|(2\spadesuit \text{ in 5 draws})) = \boxed{\frac{\binom{13}{2}\binom{39}{2}}{\binom{52}{5}}\frac{11}{47}}$

(2) [2 Pts] On average 8 calls per hours are received at the front desk of the Math department. Assuming that the number of calls per hour can be modelled as a Poisson r.v., find the probability that at any given hour, more than 6 calls are received.

SOLUTION

Let $X \sim \text{pois}(\lambda = 8)$

$P(X > 6) = 1 - P(X \leq 6) = 1 - \texttt{ppois}(6, 8) = (...)$

(3) [6 Pts] Let the joint p.d.f. of the discrete r.v. X and Y, denoted by $f(x,y)$, be given by

	x	
y	1	2
1	0.15	0.2
2	0.25	0.4

(a) Calculate the marginal densities. Are X and Y are independent?

(b) Compute the means and variances.

(c) Compute the correlation coefficient. Are X and Y positively correlated? negatively correlated? uncorrelated?

(d) Letting $Z = 2 - X + 3Y$, compute the mean and variance of Z.

SOLUTION

(a) $f_1(x) = (0.4, 0.6)$, $f_2(y) = (0.35, 0.65)$ Since $f(1,1) = 0.15$ is different from $f_1(1)f_2(1) = (0.4)(0.35)$, then X and Y are dependent.

(b) $E[X] = \sum_1^2 x f_1(x) = 0.4 + 1.2 = 1.6$

$E[Y] = \sum_1^2 y f_1(y) = 0.35 + 1.3 = 1.65$

$var(X) = \sum_1^2 x^2 f_1(x) - (1.6)^2 = 0.4 + 2.4 - (1.6)^2 = 0.24$

$var(Y) = \sum_1^2 y^2 f_1(y) - (1.65)^2 = 0.35 + 2.6 - (1.65)^2 = 0.2275$

(c) $E[XY] = \sum_{x=1}^2 \sum_{y=1}^2 xyf(x,y) = 0.15 + 0.4 + 0.5 + 1.6 = 2.65$

$cov(X,Y) = E[XY] - E[X]E[Y] = 0.01$ Thus X and Y are positively correlated

Correlation coefficient: $\rho = \frac{cov(X,Y)}{var(X)var(Y)} = 0.183$

(d) $E[Z] = 2 - E[X] + 3E[Y] = 2 - 1.6 + (3)(1.65) = 5.35$

$var(Z) = (-1)^2 var(X) + (3)^2 var(Y) + 2(-1)(3)cov(X,Y)$

$= 0.24 + (9)(0.2275) + 2(-1)(3)(0.01) = 2.2815$

(4) [4.5 Pts] Let X be a continuous r.v. with p.d.f. $f(x) = \frac{1}{2}x, \ 0 \le x \le 2$.

(a) Compute the mean and the variance of X.

(b) Compute $P(0 < X < 1)$.

(c) Compute the 50th percentile of $f(x)$.

SOLUTION

(a) $E[X] = \int_0^2 \frac{1}{2} x^2 \, dx = \frac{1}{6}x^3 \Big|_0^2 = \frac{4}{3}$

$var(X) = E[X^2] - E[X]^2 = \int_0^2 \frac{1}{2} x^3 \, dx - (\frac{4}{3})^2 = \frac{1}{8}x^4 \Big|_0^2 - (\frac{4}{3})^2 = \frac{2}{9}$

(b) $P(0 < X < 1) = \int_0^1 \frac{1}{2} x \, dx = \frac{1}{4}x^2 \Big|_0^1 = \frac{1}{4}$

(c) Want to find c such that $P(0 < X < c) = 0.5 = \frac{1}{2}$.

$P(0 < X < c) = \int_0^c \frac{1}{2} x \, dx = \frac{1}{4}x^2 \Big|_0^c = \frac{c^2}{4}$

Thus $\frac{c^2}{4} = \frac{1}{2}$ and $c = \sqrt{2}$.

(5) [5 Pts] Let X be $N(\mu = 16, \sigma^2 = 4)$. Graph its p.d.f. and compute the following quantities.

(a) Compute the probability $P(X > 10)$.

(b) Compute the probability $P(|X - 16| < 2)$.

(c) Compute c so that $P(X \leq c) = 0.95$.

SOLUTION
Let $X \sim N(\mu = 16, \sigma = 2)$
(a) $P(X > 10) = 1 - P(X \leq 10) = 1 - \texttt{pnorm}(10, 16, 2) = (...)$
(b) $P(|X - 16| < 2) = P(14 < X < 18) = \texttt{pnorm}(18, 16, 2) - \texttt{pnorm}(14, 16, 2) = (...)$
(c) c=`qnorm(0.95,16,2)` = (...)

(6) [2 Pts] Let $\overline{X}$ be the mean of a random sample of size $n = 36$ from a uniform distribution in the interval [0,6]the exponential. Approximate the probability $P(2.7 < \overline{X} < 3.1)$.
SOLUTION

$$E[X] = \frac{a+b}{2} = \frac{0+6}{2} = 3 = E[\overline{X}]$$

$$var[X] = \frac{(b-a)^2}{12} = \frac{36}{12} = 3 \quad \Rightarrow \sigma_X = \sqrt{3} \quad \Rightarrow \sigma_{\overline{X}} = \sigma_X/\sqrt{n} = 0.269$$

$$P(2.7 < \overline{X} < 3.1) = \texttt{pnorm}(3.1, 3, 0.269) - \texttt{pnorm}(2.7, 3, 0.269) = (...)$$

(7) [2 Pts] Suppose that the scores on a standardized test in mathematics taken by students from two high schools are normally distributed with distributions $N(\mu_X, \sigma^2)$ and $N(\mu_Y, \sigma^2)$, respectively, where σ^2 is unknown. A random sample of $n_X = 9$ students from the first school yielded $\overline{x} = 81.31$, $s_x^2 = 60.76$ and a random sample of $n_Y = 15$ students from the second school yielded $\overline{y} = 78.61$, $s_y^2 = 48.24$. Compute the 95% confidence interval for $\mu_X - \mu_Y$.

SOLUTION

$C.I. : (\bar{x} - \bar{y}) \pm t(\alpha/2, n_x + n_y - 2)\, s_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} = 2.7 \pm (2.074)(3.261) = 2.7 \pm 6.762$

note: $s_p^2 = \frac{s_x^2(n_x-1)+s_y^2(n_y-1)}{n_x+n_y-2} = 59.793$, hence $s_p = 7.733$

$t(\alpha/2, n_x + n_y - 2) = \texttt{qt}(1 - 0.05/2, 22) = 2.074$

REMARK

This is related to hypothesis testing, two-sided two-sample means

$H_0 : \mu_X = \mu_Y$ vs $H_1 : \mu_X \neq \mu_Y$

$$W = \frac{\bar{x} - \bar{y}}{s_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} = \frac{2.7}{3.261} = 0.828$$

p-value: 2 `pt(-0.828,22)` $= 0.417$

We can reject H_0 provided $\alpha < 0.417$

(8) [5 Pts] The desired percentage of SiO_2 in aluminous cement is 5.5. To test whether the true average percentage is 5.5 for a certain production facility, $n = 16$ independently obtained samples were analyzed and it was found that $\bar{x} = 5.25$. Suppose that the percentage of Si O_2 found in the sample is normally distributed with $\sigma = 0.3$.

(a) Does this indicate that the true average percentage differs from 5.5? Use significance level $\alpha = 0.01$

(b) Repeat the calculation in the case where σ is unknown and $s = 0.3$.

SOLUTION

We test the two-sided hypothesis testing problem

$H_0 : \mu = 5.5$ vs $H_1 : \mu \neq 5.5$

(a) Since data are normal and the variance σ^2 is known, the test statistic W below is normally distributed

Test statistics: $W = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{(4)(5.25-5.5)}{0.3} = -3.333$

For $\alpha = 0.01$, $z_{\alpha/2} = \texttt{qnorm}(1 - 0.01/2) = 2.576$.

Since $W < -z_{\alpha/2}$, then we reject H_0.

(b) Since data are normal and the variance σ^2 is unknown, the test statistic W below is t-distributed

Test statistics: $W = \frac{\bar{X}-\mu}{s/\sqrt{n}} = \frac{(4)(5.25-5.5)}{0.3} = -3.333$

For $\alpha = 0.01$, $t_{\alpha/2,n-1} = \texttt{qt}(1 - 0.01/2, 15) = 2.947$.

Since $W < -t_{\alpha/2,n-1}$, then we reject H_0.