

北京大学数学科学学院期末考试

20 21 -20 22 学年第 二 学期

考试科目: 概率统计 B 考试时间: 2022 年 6 月 23 日
姓 名: _____ 学 号: _____
本试题共 七 道大题, 满分 100 分

1. (12 分) 设二维随机变量 (X, Y) 的联合分布列如下表

$X \backslash Y$	0	1	2	3
0	0	0.01	0.01	0.01
1	0.01	0.02	0.03	0.02
2	0.03	0.04	0.05	0.04
3	0.05	0.05	0.05	0.06
4	0.07	0.06	0.05	0.06
5	0.09	0.08	0.06	0.05

试求条件期望 $E(X|Y=2)$ 和 $E(Y|X=0)$.

2. (18 分) 设从总体 $X \sim N(\mu_1, \sigma_1^2)$ 和总体 $Y \sim N(\mu_2, \sigma_2^2)$ 中分别抽取容量为 $n_1 = 10, n_2 = 15$ 的独立样本, 计算得到 $\bar{x} = 82, s_x^2 = 56.5, \bar{y} = 76, s_y^2 = 52.4$,

(1). 在 $\alpha = 0.05$ 检验水平下, 是否可以认为 σ_1^2 与 σ_2^2 显著不同.

(2). 求 σ_1^2/σ_2^2 的置信水平为 95% 的置信区间。

(3). 求 $\mu_1 - \mu_2$ 的置信水平为 95% 的置信区间。

3. (15 分) 我们测量 30 个工件的长度 (单位: cm)。测得样本均值为 9.952, 样本方差为 0.009019。已知测量值相互独立, 服从一个完全未知的正态分布。

(a) 用最大似然估计法, 估计该正态分布的均值与方差。

(b) 试求均值的置信水平为 95% 的置信区间。

4. (18 分) 某生物实验室收集了 20 只小鼠的体重与某日的进食量的数据 (单位: g)。

x (小鼠体重/g)	22.0	26.9	18.4	24.0	19.3	21.0	28.5	24.5	21.7	25.5
y (单日进食量/g)	3.66	4.18	3.37	3.75	3.36	3.86	4.39	3.97	3.67	3.62
x (小鼠体重/g)	22.6	25.0	25.0	23.4	24.3	31.7	18.2	23.5	21.0	28.3
y (单日进食量/g)	3.76	4.28	3.67	3.86	3.86	4.82	3.20	3.87	3.71	4.21

(a) 用最小二乘法, 求解回归直线 $y = a + bx$, 其中 x 为小鼠体重, y 为单日进食量。

(b) 求该回归模型中的复相关系数平方 (R^2)。

5. (12 分) 设某次考试共 100 道选择题, 每道题为 4 选 1 的单选题。设某考生一点也不懂, 完全靠蒙 (等概率地从四个选项里随机选一个), 分别估算选对至少 20 道题和至少 40 题的概率。

6. (15 分) 有一批文物共 n 件, 每种文物用两种方法测定年龄。第 i 件文物用方法一测定年龄为 x_i , 用方法二测定年龄为 y_i , ($i = 1, 2, \dots, n$). 不妨设每种测量方法给出的测量服从正态分布。

(1). 如果要检验两种测量方法是否有差异, 请给出你的统计方法?

(2). 如果 $n = 12$, 测得的数据如下,

x	384.2	325.4	331.8	339.3	393.2	362.8	338.2	338.8	387.1	370.1	336.4	336.2
y	401.2	341.5	361.7	389.2	422.2	402.8	351.3	309.2	395.2	358.3	406.5	344.4

问在显著水平 $\alpha = 0.05$ 下, 能否认为两种测量方法有显著差异?

7. (10 分) 假设未名湖的苍鹭每隔一个参数为 λ 的指数分布时间 (单位: 天) 会捉一条鱼, 每次捉鱼的时间间隔独立同分布。张老师想要了解苍鹭每次捉鱼需要的平均时间 $E[X]$, 因此委托一位同学去现场记录。但该同学因为粗心没有记录每次时间间隔, 仅仅记录了每天苍鹭捉鱼的条数 n_1, n_2, \dots, n_m . 请利用目前现有的数据, 构造一个 $E[X]$ 的强相合估计。

附录:

a. 标准正态分布上分位点: $z_{0.05} = 1.65, z_{0.025} = 1.96, z_{0.005} = 2.33$.

b. t 分布上分位点: $t_{0.05}(9) = 1.833, t_{0.025}(9) = 2.262, t_{0.05}(10) = 1.812, t_{0.025}(10) = 2.228, t_{0.05}(11) = 1.796, t_{0.025}(11) = 2.201, t_{0.05}(12) = 1.782, t_{0.025}(12) = 2.179, t_{0.05}(14) = 1.761, t_{0.025}(14) = 2.145, t_{0.05}(15) = 1.753, t_{0.025}(15) = 2.131, t_{0.05}(23) = 1.714, t_{0.025}(23) = 2.069, t_{0.05}(24) = 1.711, t_{0.025}(24) = 2.064, t_{0.05}(25) = 1.708, t_{0.025}(25) = 2.060, t_{0.05}(29) = 1.699, t_{0.025}(29) = 2.045, t_{0.05}(30) = 1.697, t_{0.025}(30) = 2.042$.

c. 卡方分布上分位点: $\chi_{0.025}^2(14) = 26.12, \chi_{0.05}^2(14) = 23.69, \chi_{0.95}^2(14) = 6.57, \chi_{0.975}^2(14) = 5.63, \chi_{0.025}^2(15) = 27.49, \chi_{0.05}^2(15) = 25.00, \chi_{0.95}^2(15) = 7.26, \chi_{0.975}^2(15) = 6.26$.

d. F 分布上分位点: $F_{0.05}(9, 14) = 2.646, F_{0.025}(9, 14) = 3.209, F_{0.05}(14, 9) = 3.025, F_{0.025}(14, 9) = 3.798, F_{0.05}(9, 15) = 2.588, F_{0.025}(9, 15) = 3.123, F_{0.05}(15, 9) = 3.006, F_{0.025}(15, 9) = 3.769, F_{0.05}(10, 14) = 2.602, F_{0.025}(10, 14) = 3.147, F_{0.05}(14, 10) = 2.865, F_{0.025}(14, 10) = 3.550, F_{0.05}(10, 15) = 2.544, F_{0.025}(10, 15) = 3.060, F_{0.05}(15, 10) = 2.845, F_{0.025}(15, 10) = 3.522$.

e. 指数分布 $\text{Exp}(\lambda): p(x) = \lambda e^{-\lambda x}, x \geq 0$.

f. 正态分布 $N(\mu, \sigma^2): p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in (-\infty, +\infty)$.