Exam in Statistical Methods, 2018-10-25

Time allowed: kl: 8-12

Allowed aids: Calculator. One handwritten A4 paper (both sides) with the students own notes.

Assisting teacher: Hector Rodriguez-Deniz

Grades: A=19-20 points, B=17-18p, C=14-16p, D=12-13p, E=10-11p, F=0-9p

Provide a detailed report that shows motivation of the results.

1

Let $f(y) = cy^2,\ 0 \leq y \leq 2$ be the density function of Y.

a) Find c so that $f(y)$ is a density function. 1p

b) Find $E(Y^2)$ and $V(Y)$. 1p

c) Find the cumulative distribution function $F(y)$, and check that $\lim_{y \to 0} F(y) = 0$ and $\lim_{y \to 2} F(y) = 1$. 1p

d) Calculate the probability $P\left(\frac{3}{4} \leq Y \leq \frac{5}{4}\right)$. 1p

2

Let the bivariate random variable (X, Y) have joint density function:

$$f(x,y) = \begin{cases} \frac{x^2 e^{-x/y}}{2y^3}, & 0 \leq x < \infty, \quad 0 \leq y \leq 1 \\ 0, & elsewhere \end{cases}$$

a) Find $f(y)$. What is the "name" of this distribution? 2p

b) Find $V(2Y + 16)$. 1p

c) Find $f(x|y)$. What is the "name" of this distribution? 1p

Hints: $\int x^2 e^{-x/a}\, dx = (-ax^2 - a^2 2x - 2a^3)e^{-x/a}$, and $\lim_{x \to \infty} \frac{ax^2+bx+c}{e^{dx}} = 0$

3

Let $Y_1, Y_2, \ldots, Y_n$ denote n independent and identically distributed random variables from the following distribution with parameters α, k,

$$f(y|\alpha, k) = \begin{cases} \frac{\alpha k^\alpha}{y^{\alpha+1}}, & k \leq y < \infty \\ 0, & elsewhere \end{cases}$$

a) Estimate k using the Method of Moments when $\alpha = 2$. 2p

b) Estimate α using Maximum Likelihood when $k = 1$. 3p

Hint: $\int \frac{1}{(x+a)^2} dx = -\frac{1}{x+a}$

4

We are interested in study the relationship between body fat and several possible predictor variables. The sample consists of 20 females with ages between 25-30 years, including the following variables:

Y = Amount of body fat (%)
X1 = Triceps skinfold thickness (mm)
X2 = Thigh circumference (cm)
X3 = Midarm circumference (cm)

We set up the following multiple linear regression model, $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$, and estimated it using ordinary least-squares (OLS), obtaining the following estimates and results:

$$SSE = 98.405$$

$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} 117.085 \\ 4.334 \\ -2.857 \\ -2.186 \end{pmatrix} \qquad (X'X)^{-1} = \begin{pmatrix} 1{,}618.867 & 48.810 & -41.849 & -25.799 \\ 48.810 & 1.479 & -1.265 & -0.779 \\ -41.849 & -1.265 & 1.084 & 0.666 \\ -25.799 & -0.779 & 0.666 & 0.414 \end{pmatrix}$$

a) Test $H_0: \beta_3 = 0$ against $H_a: \beta_3 < 0$, at 5% significance level. Interpret your result. What are your conclusions for the same test at 10% significance level? 3p

5

Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a Poisson distributed population with density $p(y) = \frac{\theta^y e^{-\theta}}{y!}, y = 0,1,2\ldots,$ and unknown parameter $0 < \theta$. We want to perform Bayesian inference on θ by defining a conjugate Gamma prior $p(\theta) \sim Gamma(\alpha, \beta)$.

a) Find the posterior distribution for θ, including the constant k. 2p
b) Find an expression for the estimator $\hat{\theta}_{Bayes}$. 1p
c) If we collect $n = 100$ samples and the sum of the counts is $\sum y = 832$, what would be your Bayes estimate for θ if you use a prior $p(\theta) \sim Gamma(2,4)$?. 1p

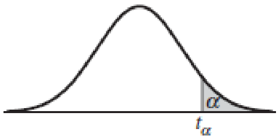
Discrete Distributions

Distribution	Probability Function	Mean	Variance
Binomial	$p(y) = \binom{n}{y} p^y (1-p)^{n-y};$ $y = 0, 1, \ldots, n$	np	$np(1-p)$
Geometric	$p(y) = p(1-p)^{y-1};$ $y = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Hypergeometric	$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}};$ $y = 0, 1, \ldots, n$ if $n \leq r$, $y = 0, 1, \ldots, r$ if $n > r$	$\frac{nr}{N}$	$n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$
Poisson	$p(y) = \frac{\lambda^y e^{-\lambda}}{y!};$ $y = 0, 1, 2, \ldots$	λ	λ
Negative binomial	$p(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r};$ $y = r, r+1, \ldots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$

Continuous Distributions

Distribution	Probability Function	Mean	Variance
Uniform	$f(y) = \frac{1}{\theta_2 - \theta_1}; \theta_1 \le y \le \theta_2$	$\frac{\theta_1 + \theta_2}{2}$	$\frac{(\theta_2 - \theta_1)^2}{12}$
Normal	$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\left(\frac{1}{2\sigma^2}\right)(y-\mu)^2\right]$ $-\infty < y < +\infty$	μ	σ^2
Exponential	$f(y) = \frac{1}{\beta} e^{-y/\beta}; \quad \beta > 0$ $0 < y < \infty$	β	β^2
Gamma	$f(y) = \left[\frac{1}{\Gamma(\alpha)\beta^\alpha}\right] y^{\alpha-1} e^{-y/\beta};$ $0 < y < \infty$	$\alpha\beta$	$\alpha\beta^2$
Chi-square	$f(y) = \frac{(y)^{(v/2)-1} e^{-y/2}}{2^{v/2}\Gamma(v/2)};$ $y^2 > 0$	v	$2v$
Beta	$f(y) = \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right] y^{\alpha-1}(1-y)^{\beta-1};$ $0 < y < 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Table 5 Percentage Points of the *t* Distributions



$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	df
3.078	6.314	12.706	31.821	63.657	1
1.886	2.920	4.303	6.965	9.925	2
1.638	2.353	3.182	4.541	5.841	3
1.533	2.132	2.776	3.747	4.604	4
1.476	2.015	2.571	3.365	4.032	5
1.440	1.943	2.447	3.143	3.707	6
1.415	1.895	2.365	2.998	3.499	7
1.397	1.860	2.306	2.896	3.355	8
1.383	1.833	2.262	2.821	3.250	9
1.372	1.812	2.228	2.764	3.169	10
1.363	1.796	2.201	2.718	3.106	11
1.356	1.782	2.179	2.681	3.055	12
1.350	1.771	2.160	2.650	3.012	13
1.345	1.761	2.145	2.624	2.977	14
1.341	1.753	2.131	2.602	2.947	15
1.337	1.746	2.120	2.583	2.921	16
1.333	1.740	2.110	2.567	2.898	17
1.330	1.734	2.101	2.552	2.878	18
1.328	1.729	2.093	2.539	2.861	19
1.325	1.725	2.086	2.528	2.845	20
1.323	1.721	2.080	2.518	2.831	21
1.321	1.717	2.074	2.508	2.819	22
1.319	1.714	2.069	2.500	2.807	23
1.318	1.711	2.064	2.492	2.797	24
1.316	1.708	2.060	2.485	2.787	25
1.315	1.706	2.056	2.479	2.779	26
1.314	1.703	2.052	2.473	2.771	27
1.313	1.701	2.048	2.467	2.763	28
1.311	1.699	2.045	2.462	2.756	29
1.282	1.645	1.960	2.326	2.576	inf.