2013 Question 1

1. (a) Suppose that the random variables X and Y are jointly distributed with probability density function

$$f(x,y) = \begin{cases} \frac{1}{3\log 2}\left(\frac{x}{y} + \frac{y}{x}\right) & 1 \le x \le 2, 1 \le y \le 2; \\ 0 & \text{otherwise} \end{cases}$$

i) Find the marginal probability density function of X; [3]

PDT Chapter 3

2. If X and Y are continuous random variables with joint density function $f_{(X,Y)}(x,y)$, then

$$F_X(x) = \mathbb{P}(X \leq x, -\infty < Y < \infty) = \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f_{(X,Y)}(u,y)dy \right) du.$$

It follows that the *marginal density function* of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)dy.$$

2013 Question 1

1. (a) Suppose that the random variables X and Y are jointly distributed with probability density function

$$f(x,y) = \begin{cases} \frac{1}{3\log 2}\left(\frac{x}{y} + \frac{y}{x}\right) & 1 \leq x \leq 2, 1 \leq y \leq 2; \\ 0 & \text{otherwise} \end{cases}$$

i) Find the marginal probability density function of X; [3]

ii) Calculate $E(XY)$; [3]

PDT Chapter 3

Lemma 3 ***Law of the unconscious statistician - 2 variables***

If $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is a sufficiently nice function then

$$\mathbb{E}[g(X,Y)] = \sum_{(x,y)} g(x,y)p_{(X,Y)}(x,y)$$

when X and Y are discrete and

$$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f_{(X,Y)}(x,y)dxdy$$

when X and Y are continuous.

2013 Question 1

1. (a) Suppose that the random variables X and Y are jointly distributed with probability density function

$$f(x,y) = \begin{cases} \frac{1}{3\log 2}\left(\frac{x}{y} + \frac{y}{x}\right) & 1 \leq x \leq 2, 1 \leq y \leq 2; \\ 0 & \text{otherwise} \end{cases}$$

i) Find the marginal probability density function of X; [3]

ii) Calculate $E(XY)$; [3]

iii) Find the conditional probability density function $f(y|x)$ for $1 \leq x \leq 2, 1 \leq y \leq 2$ and hence evaluate $P(Y < 1.5|X = 1)$. [4]

Lemma 6 ***Conditional density function.***

The conditional density function of Y *given* X *is the function* $f_{Y|X}$ *given by*

$$f_{Y|X}(y|x) = \frac{f_{(X,Y)}(x, y)}{f_X(x)}$$

for any x *such that* $f_X(x) > 0$.

Definition 9 ***Conditional distribution function.***

Let X *and* Y *be continuous random variables. The conditional distribution function of the random variable* Y *given* $X = x$ *is the function* $F_{Y|X}(\cdot|x)$ *given by*

$$F_{Y|X}(y|x) = \int_{-\infty}^{y} \frac{f_{(X,Y)}(x, v)}{f_X(x)} dv \tag{2}$$

for any x *such that* $f_X(x) > 0$.

2013 Question 1

(b) Let X and Y be independent random variables with the probability distribution functions defined as

$$F_X(x) = P(X \leq x) \quad \textit{and} \quad F_Y(y) = P(Y \leq y).$$

i) Let $U = \min\{X, Y\}$ and $V = \max\{X, Y\}$. Show that

$$\begin{aligned} F_U(u) &= 1 - \{1 - F_X(u)\}\{1 - F_Y(u)\} \\ F_V(v) &= F_X(v)F_Y(v) \end{aligned}$$

[4]

Definition 4 ***Independence.*** *X and Y are independent if $\{X \leq x\}$ and $\{Y \leq y\}$ are independent events for all $x, y \in \mathbb{R}$. In other words, X and Y are independent if*

$$F_{(X,Y)}(x, y) = F_X(x)F_Y(y).$$

2013 Question 1

(b) Let X and Y be independent random variables with the probability distribution functions defined as

$$F_X(x) = P(X \leq x) \quad and \quad F_Y(y) = P(Y \leq y).$$

i) Let $U = \min\{X, Y\}$ and $V = \max\{X, Y\}$. Show that

$$\begin{aligned} F_U(u) &= 1 - \{1 - F_X(u)\}\{1 - F_Y(u)\} \\ F_V(v) &= F_X(v)F_Y(v) \end{aligned}$$

[4]

Now let X and Y be the independent exponential random variables with density functions

$$f(x) = e^{-x} \quad (x \geq 0) \qquad and \qquad f(y) = e^{-y} \quad (y \geq 0).$$

ii) Find the distribution function of $V = \max\{X, Y\}$. [2]

2013 Question 1

(b) Let X and Y be independent random variables with the probability distribution functions defined as

$$F_X(x) = P(X \leq x) \quad \text{and} \quad F_Y(y) = P(Y \leq y).$$

i) Let $U = \min\{X, Y\}$ and $V = \max\{X, Y\}$. Show that

$$\begin{aligned} F_U(u) &= 1 - \{1 - F_X(u)\}\{1 - F_Y(u)\} \\ F_V(v) &= F_X(v)F_Y(v) \end{aligned}$$

[4]

Now let X and Y be the independent exponential random variables with density functions

$$f(x) = e^{-x} \quad (x \geq 0) \qquad \text{and} \qquad f(y) = e^{-y} \quad (y \geq 0).$$

ii) Find the distribution function of $V = \max\{X, Y\}$. [2]

iii) Show that $Z = X + \frac{1}{2}Y$ has the same distribution function as V. [4]

PDT Chapter 4

Theorem 1 *If X and Y have joint density function $f_{(X,Y)}(x,y)(\cdot,\cdot)$, then $Z = X+Y$ has density function*

$$f_Z(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(x, z-x)dx.$$

Moreover, if they are independent we have

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)dy$$

in which case, the function f_Z, also denoted by f_{X+Y}, is called the convolution of f_X and f_Y and is written as

$$f_{X+Y} = f_X * f_Y.$$

Proof.

$$F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(X+Y \leq z) \quad = \quad \int\int_{x+y\leq z} f_{(X,Y)}(x,y)dxdy$$

2012 Question 2

2. The random variable Z has the χ^2_k distribution (k is the degrees of freedom) which has the moment generating function (mgf)

$$M_Z(\theta) = (1 - 2\theta)^{-\frac{k}{2}} \quad \text{for} \quad \theta < \frac{1}{2}.$$

   (a) Using the mgf, find the mean and variance of Z. [4]

PDT Chapter 4, page 13 & Chapter 2, Lemma 5

where μ_k is the kth moment of X, i.e, $\mu_r = \mathbb{E}[X^r]$ for $r = 1, 2, \ldots$. We recover the moments by the operation of differentiation, i.e,

$$\mu_k = \frac{\partial^k}{\partial \theta^k} M_X(\theta)|_{\theta=0}.$$

5. $\mathrm{var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

2012 Question 2

2. The random variable Z has the χ^2_k distribution (k is the degrees of freedom) which has the moment generating function (mgf)

$$M_Z(\theta) = (1 - 2\theta)^{-\frac{k}{2}} \quad \text{for} \quad \theta < \frac{1}{2}.$$

(a) Using the mgf, find the mean and variance of Z. [4]

(b) Suppose X has standard normal distribution $N(0, 1)$. Using integration, show that $Y = X^2$ has the χ^2_1 distribution. [6]

PDT Chapter 4, pages 11 & 12

Remark: If X has the generating function $M_X(\theta)$, then the generating function of $Y = g(X)$ can be found in the following way:

$$M_Y(\theta) = \mathbb{E}[e^{\theta Y}] = \mathbb{E}[e^{\theta g(X)}],$$

Uniqueness

If X and Y have the same moment generating function, then X and Y have the same distribution.

2012 Question 2

2. The random variable Z has the χ^2_k distribution (k is the degrees of freedom) which has the moment generating function (mgf)

$$M_Z(\theta) = (1-2\theta)^{-\frac{k}{2}} \quad \text{for} \quad \theta < \frac{1}{2}.$$

   (a) Using the mgf, find the mean and variance of Z. [4]

   (b) Suppose X has standard normal distribution $N(0,1)$. Using integration, show that $Y = X^2$ has the χ^2_1 distribution. [6]

   (c) Show that if $Y_1, Y_2, \cdots, Y_n$ are independent, each with a χ^2_1 distribution, then $V = \sum_{i=1}^n Y_i$ has a χ^2_n distribution. [2]

Independence

If X and Y are independent random variables then

$$M_{X+Y}(\theta) = M_X(\theta)M_Y(\theta)$$

for all θ for which both $M_X(\theta)$ and $M_Y(\theta)$ are defined.

2012 Question 2

2. The random variable Z has the χ^2_k distribution (k is the degrees of freedom) which has the moment generating function (mgf)

$$M_Z(\theta) = (1-2\theta)^{-\frac{k}{2}} \quad \text{for} \quad \theta < \frac{1}{2}.$$

(a) Using the mgf, find the mean and variance of Z. [4]

(b) Suppose X has standard normal distribution $N(0,1)$. Using integration, show that $Y = X^2$ has the χ^2_1 distribution. [6]

(c) Show that if $Y_1, Y_2, \cdots, Y_n$ are independent, each with a χ^2_1 distribution, then $V = \sum_{i=1}^{n} Y_i$ has a χ^2_n distribution. [2]

(d) Use the previous results and the central limit theorem to find the approximate probability that $V < 310$ when $n = 300$. [4]

PDT Chapter 5

Theorem 3 (central limit theorem)**.** Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with finite mean μ and finite non-zero variance σ^2, and let $S_n = X_1 + \cdots + X_n$. Then

$$\frac{\frac{S_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} N(0,1) \qquad \text{as } n \longrightarrow \infty.$$

2012 Question 2

(e) Suppose that $X_1, \cdots, X_n$ is a random sample drawn from a normal distribution $N(\mu, \sigma^2)$. Let

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i; \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

State without proof the distributions of the two statistics $\overline{X}$ and

$$\frac{(n-1)S^2}{\sigma^2}.$$

Further, find the distribution of

$$T = \frac{\sqrt{n}(\overline{X} - \mu)}{S}$$

by quoting a proper definition. Here S is the square root of S^2. [Note: all the parameters in these distributions should be properly stated.] [4]

PDT Chapter 5, section 5.4

If $X_1, X_2, \cdots$ are independent $N(\mu, \sigma^2)$ variables then

$$\overline{X} \sim N(\mu, \sigma^2/n),$$

Using (2), it can be shown that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

If $Z \sim N(0,1)$ and $Y \sim \chi^2_r$ *independently* of each other then

$$T = \frac{Z}{(Y/r)^{1/2}}$$

is defined to have a t distribution with r degrees of freedom.

2012 Question 3

3. For a productive pair from a particular species of bird, the number X of eggs laid per season has the probability mass function

$$P(X = k) = C\frac{e^{-\lambda}\lambda^k}{k!} \quad \text{for} \quad k = 1, 2, 3, \cdots$$

Suppose $X_1, \cdots, X_n$ is a random sample drawn from the population.

(a) Find the constant C. [4]

2012 Question 3

3. For a productive pair from a particular species of bird, the number X of eggs laid per season has the probability mass function

$$P(X = k) = C\frac{e^{-\lambda}\lambda^k}{k!} \quad \text{for} \quad k = 1, 2, 3, \cdots$$

Suppose $X_1, \cdots, X_n$ is a random sample drawn from the population.

(a) Find the constant C. [4]

(b) Find an equation for determining $\widetilde{\lambda}$, the estimator of λ by Method of Moments based on $X_1, \cdots, X_n$. (Note: Do not attempt to solve the equation.) [4]

Let $\mathbf{X} \equiv (X_1, X_2, \ldots, X_n)$ be a random sample of size n from some family of distributions depending upon the vector of q parameters, $\theta = (\theta_1, \theta_2, \ldots, \theta_q)$. Assume that the first q moments $\mu_j,\ j = 1, \ldots, q$ of the distributions exist,

$$\mu_j(\theta_1, \theta_2, \ldots, \theta_q) = E[X_i^j; \theta_1, \theta_2, \ldots, \theta_q] \qquad j = 1, \ldots, q.$$

Let $m_j,\ j = 1, \ldots, q$ be the corresponding sample moments,

$$m_j = \frac{1}{n}\sum_{i=1}^{n} X_i^j \qquad j = 1, \ldots, q.$$

According to the method of moments, we match the sample and population/distribution moments. We solve the following set of q simultaneous equations for the q unknown parameter values,

$$\begin{aligned} m_1 &= \mu_1(\hat\theta_1, \hat\theta_2, \ldots, \hat\theta_q) \\ m_2 &= \mu_2(\hat\theta_1, \hat\theta_2, \ldots, \hat\theta_q) \\ &\vdots \\ m_q &= \mu_q(\hat\theta_1, \hat\theta_2, \ldots, \hat\theta_q), \end{aligned} \tag{1}$$

to obtain the estimator $(\hat\theta_1, \hat\theta_2, \ldots, \hat\theta_q)$.

2012 Question 3

3. For a productive pair from a particular species of bird, the number X of eggs laid per season has the probability mass function

$$P(X = k) = C\frac{e^{-\lambda}\lambda^k}{k!} \quad \text{for} \quad k = 1, 2, 3, \cdots$$

Suppose $X_1, \cdots, X_n$ is a random sample drawn from the population.

(a) Find the constant C. [4]

(b) Find an equation for determining $\widetilde{\lambda}$, the estimator of λ by Method of Moments based on $X_1, \cdots, X_n$. (Note: Do not attempt to solve the equation.) [4]

(c) Show that the maximum likelihood estimator $\widehat{\lambda}$ of λ based on $X_1, \cdots, X_n$ is identical with the estimator of the method of moments. (Note: Do not attempt to solve the equation.) [6]

Inference Chapter 2, pages 3 & 4

Definition

For each $\mathbf{x} \in \mathcal{X}$, let $\hat{\theta}(\mathbf{x})$ be such that, with $\mathbf{x}$ held fixed, the likelihood function $\{L(\theta; \mathbf{x}) : \theta \in \Theta\}$ attains its maximum value as a function of θ at $\hat{\theta}(\mathbf{x})$. The estimator $\hat{\theta}(\mathbf{X})$ is then said to be a *maximum likelihood estimator* (MLE) of θ.

Assuming that the likelihood function $L(\theta; \mathbf{x})$ is a continuously differentiable function of θ, given $\mathbf{x} \in \mathcal{X}$, an interior stationary point of $\ln L(\theta; \mathbf{x})$ or of $L(\theta; \mathbf{x})$ is given by a solution of the *likelihood equations*,

$$\frac{\partial \ln L(\theta; \mathbf{x})}{\partial \theta_j} = 0, \qquad j = 1, \ldots, q. \tag{4}$$

A solution of Equations (4) may or may not be unique, and may or may not give us a MLE. In many of the standard cases, solution of the Equations (4) does give us the MLE, but we cannot take this for granted.

2012 Question 3

3. For a productive pair from a particular species of bird, the number X of eggs laid per season has the probability mass function

$$P(X=k) = C\frac{e^{-\lambda}\lambda^k}{k!} \quad \text{for} \quad k = 1, 2, 3, \cdots$$

Suppose $X_1, \cdots, X_n$ is a random sample drawn from the population.

(a) Find the constant C. [4]

(b) Find an equation for determining $\widetilde{\lambda}$, the estimator of λ by Method of Moments based on $X_1, \cdots, X_n$. (Note: Do not attempt to solve the equation.) [4]

(c) Show that the maximum likelihood estimator $\widehat{\lambda}$ of λ based on $X_1, \cdots, X_n$ is identical with the estimator of the method of moments. (Note: Do not attempt to solve the equation.) [6]

(d) Find the Fisher information of λ from the simple random sample $X_1, \cdots, X_n$. [4]

Inference Chapter 3

where $I(\theta)$ *is the* Fisher information,

$$I(\theta) = E_\theta\left[\left(\frac{\partial \ln f(\mathbf{X};\theta)}{\partial\theta}\right)^2\right] \qquad \theta \in \Theta. \tag{2}$$

Theorem 3 *Under appropriate regularity conditions,*

$$I(\theta) = E_\theta\left[-\frac{\partial^2 \ln f(\mathbf{X};\theta)}{\partial\theta^2}\right] \qquad \theta \in \Theta.$$

2012 Question 3

3. For a productive pair from a particular species of bird, the number X of eggs laid per season has the probability mass function

$$P(X = k) = C\frac{e^{-\lambda}\lambda^k}{k!} \quad \text{for} \quad k = 1, 2, 3, \cdots$$

Suppose $X_1, \cdots, X_n$ is a random sample drawn from the population.

(a) Find the constant C. [4]

(b) Find an equation for determining $\widetilde{\lambda}$, the estimator of λ by Method of Moments based on $X_1, \cdots, X_n$. (Note: Do not attempt to solve the equation.) [4]

(c) Show that the maximum likelihood estimator $\widehat{\lambda}$ of λ based on $X_1, \cdots, X_n$ is identical with the estimator of the method of moments. (Note: Do not attempt to solve the equation.) [6]

(d) Find the Fisher information of λ from the simple random sample $X_1, \cdots, X_n$. [4]

(e) Quoting any appropriate asymptotic properties of the maximum likelihood estimators, deduce the approximate distribution of $\widehat{\lambda}$ for large n. [2]

It can be shown, using the Central Limit theorem, that MLEs are asymptotically normally distributed and *asymptotically efficient*, i.e., in the case of $\Theta \subseteq \mathbb{R}$, for all $\theta \in \Theta$

$$\sqrt{n}(\hat{\tau}_n - \tau(\theta)) \to N\left(0, \frac{\left(\frac{d\tau}{d\theta}\right)^2}{i(\theta)}\right).$$

Put less formally, for large n,

$$\hat{\tau}_n \sim N\left(\tau(\theta), \frac{\left(\frac{d\tau}{d\theta}\right)^2}{I(\theta)}\right)$$

2013 Question 4

4. (a) Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(0, \sigma^2)$.

i) Find the most powerful test for the simple hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 = \sigma_1^2$ where $\sigma_1^2 > \sigma_0^2$. [3]

Theorem 1 (The Neyman-Pearson Lemma) *Given the simple hypotheses H_0 and H_1, let $\mathcal{C}$ be the critical region as specified in Equation (1) of a likelihood ratio test for some given $k > 0$. Let α be the significance level and β the power of this test. Any other test with significance level less than or equal to α has power less than or equal to β.*

$$\mathcal{C} = \{\mathbf{x} \in \mathcal{X} : \frac{f_1(\mathbf{x})}{f_0(\mathbf{x})} \geq k\} \tag{1}$$

for some positive constant k. This is known as a *likelihood ratio test.*

In the case of simple hypotheses, the significance level and power of a test with critical region $\mathcal{C}$ may be defined particularly simply. The *significance level* α is given by

$$\alpha = \mathbb{P}(\mathbf{X} \in \mathcal{C}; \theta_0) = \mathbb{P}(H_0 \text{ is rejected}; H_0)$$

and the *power* β by

$$\beta = \mathbb{P}(\mathbf{X} \in \mathcal{C}; \theta_1) = \mathbb{P}(H_0 \text{ is rejected}; H_1).$$

2013 Question 4

4. (a) Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(0, \sigma^2)$.

   i) Find the most powerful test for the simple hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 = \sigma_1^2$ where $\sigma_1^2 > \sigma_0^2$. [3]

   ii) Find the uniformly most powerful test for the simple hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against the composite hypothesis $H_1 : \sigma^2 > \sigma_0^2$. [3]

test is said to be a *uniformly most powerful (UMP)* test of significance level α if it has significance level less than or equal to α and $\beta(\theta) \geq \beta^*(\theta)$ for all $\theta \in \Theta_1$ for the power function $\beta^*(\theta)$ of every other test with significance level less than or equal to α.

The form of the likelihood ratio test is the same for all pairs of parameter values θ_0 and θ_1 with $\theta_1 > \theta_0$. It follows that a test with critical region of the form (6) is UMP for testing

$$H_0 : \theta = \theta_0 \qquad \text{against} \qquad H_1 : \theta > \theta_0,$$

or, indeed, for testing

$$H_0 : \theta \leq \theta_0 \qquad \text{against} \qquad H_1 : \theta > \theta_0.$$

2013 Question 4

4. (a) Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(0, \sigma^2)$.

   i) Find the most powerful test for the simple hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 = \sigma_1^2$ where $\sigma_1^2 > \sigma_0^2$. [3]

   ii) Find the uniformly most powerful test for the simple hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against the composite hypothesis $H_1 : \sigma^2 > \sigma_0^2$. [3]

   iii) Let $\alpha = P(\sum_{i=1}^{n} X_i^2 > c|\sigma_0^2)$ where c is an unknown parameter to be determined. Given $\sigma_0^2 = 4$ and $n = 15$, find the value of c such that the significance level $\alpha = 0.05$. [3]

Statistical tables

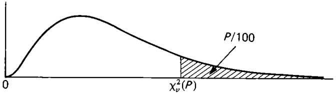
TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.



(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72

2013 Question 4

4. (a) Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(0, \sigma^2)$.

   i) Find the most powerful test for the simple hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 = \sigma_1^2$ where $\sigma_1^2 > \sigma_0^2$. [3]

   ii) Find the uniformly most powerful test for the simple hypothesis $H_0 : \sigma^2 = \sigma_0^2$ against the composite hypothesis $H_1 : \sigma^2 > \sigma_0^2$. [3]

   iii) Let $\alpha = P(\sum_{i=1}^{n} X_i^2 > c|\sigma_0^2)$ where c is an unknown parameter to be determined. Given $\sigma_0^2 = 4$ and $n = 15$, find the value of c such that the significance level $\alpha = 0.05$. [3]

   iv) Given $n = 15$ and c is the value found in part (iii), find the approximate value of $\beta = P(\sum_{i=1}^{n} X_i^2 < c|\sigma_1^2 = 16)$. [3]

Statistical tables

TABLE 7. THE χ^2-DISTRIBUTION FUNCTION

$\nu =$	15	16	17	18	19	20	21	22	23
$x = 3$	0·0004	0·0002	0·0001						
4	·0023	·0011	·0005	0·0002	0·0001				
5	0·0079	0·0042	0·0022	0·0011	0·0006	0·0003	0·0001	0·0001	
6	·0203	·0119	·0068	·0038	·0021	·0011	·0006	·0003	0·0001
7	·0424	·0267	·0165	·0099	·0058	·0033	·0019	·0010	·0005
8	·0762	·0511	·0335	·0214	·0133	·0081	·0049	·0028	·0016
9	·1225	·0866	·0597	·0403	·0265	·0171	·0108	·0067	·0040

2013 Question 4

(b) The probability density function of gamma distribution with parameters $\alpha > 0$ and $\theta > 0$ is given by

$$f(x) = \frac{x^{\alpha-1}e^{-x/\theta}}{\Gamma(\alpha)\theta^{\alpha}} \quad \text{for} \quad x > 0.$$

Let $X_1, X_2, \cdots, X_n$ be a random sample from a gamma distribution with the parameter $\alpha = 1$ and the unknown parameter θ. Let $\tau = \log\theta$. Suppose that the prior density function of τ is an improper uniform prior, i.e,

$$\pi(\tau) \propto 1, \quad \tau \in (-\infty, \infty).$$

i) Find the prior density function of θ up to a constant of proportionality. [2]

Note also that if the parameter space Θ is not bounded then there does not exist a uniform distribution on Θ. If we take $\pi(\theta) = k$ for some constant $k \geq 0$ then

$$\int_{\Theta} \pi(\theta) d\theta = \begin{cases} \infty & k > 0 \\ 0 & k = 0 \end{cases}$$

Thus it is not possible to make the integral take the value 1. However, in view of Equation (2), we only have to specify $\pi(\theta)$ up to a constant of proportionality. So we can proceed, using $\pi(\theta) = k$, despite the fact that $\pi(\theta)$ does not specify a proper distribution. In such a case, $\pi(\theta)$ is known as an *improper prior*.

Example 2 *On many occasions we will be interested in calculating the distribution function of a random variable* $Y = g(X)$ *where* $g : \mathbb{R} \longrightarrow \mathbb{R}$ *is a sufficiently nice function. Let* X *be a random variable with probability density function* $f_X(\cdot)$.

So we have the following result: if g *is a monotone function (increasing or decreasing), the density function of* Y *is*

$$f_Y(y) = f_X(g^{-1}(y)) \left| \left(g^{-1}(y)\right)' \right| .$$

2013 Question 4

(b) The probability density function of gamma distribution with parameters $\alpha > 0$ and $\theta > 0$ is given by

$$f(x) = \frac{x^{\alpha-1}e^{-x/\theta}}{\Gamma(\alpha)\theta^{\alpha}} \quad \text{for} \quad x > 0.$$

Let $X_1, X_2, \cdots, X_n$ be a random sample from a gamma distribution with the parameter $\alpha = 1$ and the unknown parameter θ. Let $\tau = \log\theta$. Suppose that the prior density function of τ is an improper uniform prior, i.e,

$$\pi(\tau) \propto 1, \quad \tau \in (-\infty, \infty).$$

i) Find the prior density function of θ up to a constant of proportionality. [2]

ii) Find the posterior density function of θ up to a constant of proportionality. [2]

Inference Chapter 5

Given $\mathbf{x} \in \mathcal{X}$, the posterior density for θ is computed using Bayes' Theorem,

$$\pi(\theta|\mathbf{x}) = \frac{\pi(\theta) f(\mathbf{x}|\theta)}{m(\mathbf{x})} \qquad \theta \in \Theta, \tag{1}$$

where $m(\mathbf{x})$ is the marginal p.d.f. of $\mathbf{X}$,

$$m(\mathbf{x}) = \int_{\Theta} \pi(\theta) f(\mathbf{x}|\theta) d\theta.$$

2013 Question 4

(b) The probability density function of gamma distribution with parameters $\alpha > 0$ and $\theta > 0$ is given by

$$f(x) = \frac{x^{\alpha-1}e^{-x/\theta}}{\Gamma(\alpha)\theta^{\alpha}} \quad \text{for} \quad x > 0.$$

Let $X_1, X_2, \cdots, X_n$ be a random sample from a gamma distribution with the parameter $\alpha = 1$ and the unknown parameter θ. Let $\tau = \log\theta$. Suppose that the prior density function of τ is an improper uniform prior, i.e,

$$\pi(\tau) \propto 1, \quad \tau \in (-\infty, \infty).$$

i) Find the prior density function of θ up to a constant of proportionality. [2]

ii) Find the posterior density function of θ up to a constant of proportionality. [2]

iii) Let $z = 1/\theta$ and show that the posterior distribution of Z is a gamma distribution with parameter $\alpha = n$ and $\theta = 1/y$, where $y = \sum_{i=1}^{n} x_i$. [4]

PDT Chapter 4, pages 2–3

Example 2 *On many occasions we will be interested in calculating the distribution function of a random variable $Y = g(X)$ where $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a sufficiently nice function. Let X be a random variable with probability density function $f_X(\cdot)$.*

So we have the following result: if g is a monotone function (increasing or decreasing), the density function of Y is

$$f_Y(y) = f_X(g^{-1}(y)) \left| \left(g^{-1}(y)\right)' \right| .$$