1. 随机变量Z表示股票价格的走势，$Z=1$表示股票价格将上涨，$Z=0$则表示股价将下跌。写出Z的概率函数。
2. 随机变量的概率分布和矩有什么关系？
3. 随机变量的方差为 0 意味着什么？
4. 设随机变量ξ表示一项投资的未来收益率。ξ的概率分布的偏度和峰度分别为λ和κ。（1）说明λ的正负和大小对投资决策的影响；

   （2）说明κ的大小对投资决策的影响。
5. 用相关系刻画随机变量之间的相关性有什么缺陷？
6. ξ与η的方差相等。设$X=\xi+\eta, Y=\xi-\eta$，证明X,Y不相关。这一结论说明了什么？
7. 大数定律告诉我们了什么？
8. 为什么说样本是随机变量？
9. 为什么需要对参数进行假设检验？
10. $X \sim N(0,1)$

    （1）查附表 1，求概率$P(|X|>1.96)$。

    （4）查附表 1，给定$\alpha=0.05$，求u_α使得$P(X>u_\alpha)=\alpha$。
11. 设$\xi \sim \chi^2(n)$，依据χ^2分布的定义和标准正态分布的性质证明：$\mathrm{E}\xi=n, \mathrm{Var}(\xi)=2n$。
12. 根据 *t*-分布和 *F*-分布的定义证明：如果$\xi \sim t(n)$，则$\xi^2 \sim F(1,n)$。
13. 设$X_1, X_2, \cdots, X_n$是从总体X中抽取的样本。X服从两点分布：$P(X=1)=p$。

    （1）用矩估计法求参数p的估计；

    （2）用极大似然法求参数p的估计。
14. 为了减少吸烟对人体的危害，某地区地方政府通过提高烟草税的方法减少烟草消费。将某地区烟民月抽烟量（包）X作为总体，总体服从正态分布$X \sim N(\mu,\sigma^2)$。提高税收之前$\mu=10$。提高烟草税之后，从该地区随机抽取 101 名烟民进行月度烟草消费调查，得出样本值$X_1, X_2, \cdots, X_{101}$，为简单，假定$X_1, X_2, \cdots, X_{101}$独立同分布。经计算得出样本均值$\overline{X}=8.5$和样本方差$S^2=81$。采用假设检验的方法说明提高烟草税是否降低了烟民的烟草消费。

    （1）给出检验的原假设和备择假设；

    （2）设计检验统计量，计算其概率分布；

    （3）计算统计量值；

    （4）给出检验结果。
15. 股票 A 的日收盘价ξ为随机变量。今有两个月 40 个交易日的收盘价数据$\xi_1, \xi_2, \cdots, \xi_{40}$，计算出样本矩为

$$S^2=\frac{1}{40}\sum_{i=1}^{40}(X_i-\overline{X})^3=1.76,\quad \frac{1}{40}\sum_{i=1}^{40}(X_i-\overline{X})^4=873 \quad ,$$

    （1）根据（2.9）式计算ξ分布的偏度和峰度估计$\hat{\lambda}$和$\hat{\kappa}$；

    （2）股票收盘价服从正态分布吗？
16. 表 2.1 是某个国家 1959 年 1 月~1996 年 4 月的月度宏观经济数据。

表 2.1

IP	M2	PW	R
36	286.7	31.7	2.837

36.7	287.7	31.7	2.712
…	…	…	…
123.4	3724.4	126.4	4.96
124.5	3728	127.5	4.99

数据来源：平迪克、鲁宾费尔德

，其中 IP 为工业生产指数，M2 为广义货币，PW 为商品生产者价格指数，R 为三个月国债利率。

（1）将表 2.1 导入 EViews；

（2）将 IP、M2、PW 和 R 以组打开为数据表格；

（3）分别在同一坐标系内和分别单个画出 IP、PW 和 R 的线图（Line&Symbol）；

（4）生成新变量广义货币增长率 GM2=(M2-M2(-1))/M2(-1)；

（5）在同一坐标系内画出 IP 和 PW 的线图（Line&Symbol）和区域带状图（Area Band）；

（6）以 IP 为横轴，PW 为纵轴画出 XY-line 线图；

（7）用函数@trend 生成序列 No，取值为 1,2,3， ……

◆ 参考答案

1. Z 服从 0-1 分布，是 $n=1$ 时的二项分布，其概率函数为 $P(Z=i)=p^i(1-p)^{1-i}, i=0,1$。其中 p 为 $Z=1$ 的概率。
2. 概率分布决定矩，从概率分布可以计算出随机变量的任意矩。矩是随机变量或者其函数取值的概率平均值，是概率分布的总体特征。矩不能确定概率分布，不同的概率分布可以具有相同的矩。
3. 随机变量的方差为 0 意味着随机变量没有了随机性，等价于一个数。可以证明，如果随机变量 ξ 的方差为零 $\mathrm{Var}(\xi)=0$，数学期望为 $\mathrm{E}\xi=\mu$，则 $P(\xi=\mu)=1$。此时称 ξ 服从单点分布。
4. (1) $\lambda>0$ ($\lambda<0$) 表明分布为右（左）偏的偏态分布，即未来收益大（小）于平均收益的概率要大于小于平均收益的概率，投资项目的盈利潜力比数学期望给出的平均收益大（小）。

   （2）κ 越大（小），表明极端事件发生的概率越大。极端事件分为两种，一种是损失极端事件，一种是盈利极端事件。损失极端事件发生的概率大，表明投资风险较大，需要在投资前充分重视，采用必要的预防措施应对可能出现的极端事件。盈利极端事件发生的概率大，表明投资出现“意外”收益的可能性大，适合愿意承受较大风险而获取高额收益的投资者。
5. 相关系数只能刻画随机变量本身之间的相关性，不能刻画随机变量函数之间的相关性。例如随机变量 ξ 和 η 之间的相关系数为 0（不相关），不能得出 ξ^2 和 η^2 的相关系数为 0。
6. 证明：根据协方差的性质得出

$$\begin{aligned}\mathrm{Cov}(\xi+\eta,\xi-\eta)&=\mathrm{Cov}(\xi,\xi)-\mathrm{Cov}(\xi,\eta)+\mathrm{Cov}(\eta,\xi)-\mathrm{Cov}(\eta,\eta)\\&=\mathrm{Var}(\xi)-\mathrm{Var}(\eta)=0\end{aligned}$$

7. 大数定律表明，对一组随机变量取平均值可以有效减少随机变量的随机性，并且较为宽松的条件下，当参与平均的随机变量个数无限增加时，平均值的随机性将减少为 0，平均值以概率趋于常数。

8. 统计研究着眼于事前（ex ante）研究，即在实际抽样之前就要给出数据处理的方法—统计量，而实际抽样之前的样本是理论上的，可以是任何一个结果，因此是随机变量。进行实际抽样得到样本值之后，只需要将数值带入事先给出的统计量之中就可以得出数据处理结果。
9. 参数估计给出的结果只是真实参数的近似。由于参数估计是随机变量，而真实参数是确定的数值，因此这种近似是一种随机近似。对真实参数是否满足某个条件（例如数学期望是否为 0）的推断，只能通过其参数估计来进行。由于是用随机变量推断一个数值，推断的准确性不可能达到 100%，而只能在一定的可信度（置信水平）下进行。
10. （1）由于正态分布表只给出大于 0 的 x 对应的概率 $P(X \le x) = \Phi(x)$，而小于 0 的 x 对应的概率需要通过正态分布性质 $\Phi(x) = 1 - \Phi(-x)$ 得出。因此

$$
\begin{aligned}
P(|X| > 1.96) &= P(1.96 < X) + P(X < -1.96)) \\
&= 1 - \Phi(1.96) + \Phi(-1.96) \\
&= 2[1 - \Phi(1.96)]
\end{aligned}
$$

查附表 1 得出，$\Phi(1.96) = 0.975$，因此 $P(|X| > 1.96) = 0.05$。
11. 根据 $\chi^2(n)$ 的定义，$\xi = x_1^2 + x_2^2 + \cdots + x_n^2$，其中 $x_i \sim N(0,1), i = 1,2,\cdots,n$ 相互独立。因此

$$
\begin{aligned}
&\mathrm{E}\xi = n\mathrm{E}x_1^2 = n\mathrm{Var}(x_1) = n \\
&\mathrm{Var}(\xi) = n\mathrm{Var}(x_1^2) = n(\mathrm{E}(x_1^4) - (\mathrm{E}x_1^2)^2) = n(3-1) = 2
\end{aligned}
$$

其中 $\mathrm{E}(x_1^4)$ 为正态分布的峰度，等于 3.
12. 根据 t-分布的定义， $t = \xi / \sqrt{\eta / n}$，其中 $\xi \sim N(0,1)$，$\eta \sim \chi^2(n)$，因此

$$
t^2 = \frac{\xi^2 / 1}{\eta / n}
$$

。根据 χ^2 分布的定义，$\xi^2 \sim \chi^2(1)$。在根据 F-分布的定义得出 $t^2 \sim F(1,n)$
13. （1）根据二项分布的性质得出 $\mathrm{E}X = p$。因此得出总体矩条件 $\mathrm{E}(X - p) = 0$。给定样本 $X_1, X_2, \cdots, X_n$，根据类比原则得出样本矩条件 $n^{-1}\sum_{i=1}^{n}(X_i - \hat{\mu}) = 0$，据此得出 p 的矩估计为 $\hat{p}_{MM} = \overline{X}$。实际上也可以用样本二阶矩得出的矩条件求出 p 的矩估计：$\mathrm{Var}(X) = p(1-p)$，因此 $\mathrm{E}X^2 = \mathrm{Var}(X) + (EX)^2 = p(1-p) + p^2 = p$，得出矩条件 $\mathrm{E}(X^2 - p) = 0$，并采用类比原则得出估计量为 $\hat{p}_{MM} = n^{-1}\sum_i X_i^2$。由于 X_i 服从两点分布，取值为 0 和 1，因此 $X_i^2 = X_i$，两种方法得出的估计量相同。

（2）X_i 的概率函数为 $p^{X_i}(1-p)^{1-X_i}, i = 1,2,\cdots,n$，由此得出似然函数和对数似然函数

$$
L(X_1, X_2, \cdots, X_n; p) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i}
$$

$$
l(X_1, X_2, \cdots, X_n; p) = \ln L = \sum_{i=1}^{n} X_i \ln p + \sum_{i=1}^{n}(1 - X_i)\ln(1-p)
$$

将对数似然函数对 p 求导并令导数等于 0，得出

$$
\frac{dl(X_1, X_2, \cdots, X_n; p)}{dp} = \frac{1}{p}\sum_{i=1}^{n} X_i - \frac{1}{1-p}\sum_{i=1}^{n}(1 - X_i) = 0
$$

。解方程得出 $\hat{p}_{ML} = \overline{X}$

14. （1）根据题意，检验的备择假设应该是单侧的，即提高烟草税降低了烟民的烟草消费。因此检验的原假设和备择假设为 $\mathrm{H}_0:\ \mu = 10;\ \mathrm{H}_1: \mu < 10$。

（2）由于总体的方差未知，对均值的检验需要采用 t 检验。在原假设下 $(101-1)S^2/\sigma^2 \sim \chi^2(100)$，$\overline{X} \sim N(10, \sigma^2)$，因此 $(\overline{X}-10)/\sigma^2 \sim N(0,1)$。由此得出检验统计量

$$t = \frac{(\overline{X}-10)/\sigma}{\sqrt{\frac{(101-1)S^2/\sigma^2}{100}}} = \frac{\overline{X}-10}{S} \sim t(100)$$

（3）统计量值 $t = (8.5-10)/9 = -0.17$

（4）由于检验的备择假设为单侧，采用 $P(t < t_\alpha) = \alpha$ 确定检验的临界值 t_a 和拒绝域 $(-\infty, t_\alpha)$。取 $\alpha = 0.05$，查 t 分布表（附表 3）$t_a = -1.663$。由于统计量值没有落入拒绝域，不能拒绝原假设。即没有证据表明，提高烟草税抑制了烟草消费。

15. （1）$\hat{\lambda} = 0.0482,\ \hat{\kappa} = 7.215$；

（2）由于计算出的峰度估计值 $\hat{\kappa} = 7.215$ 远大于正态分布的峰度 3，因此股票价格不服从正态分布。

16. 略。