統計学練習問題集

西山茂、寺坂崇宏、劉慶豊
小樽商科大学

平成 24 年 5 月 30 日

データの要約

問題 1 以下のデータを質的データと量的データに分類しなさい。性別、年齢、製品カラー、売り上げ、所得税、国籍、GDP、所属。

問題 2 以下の度数分布表を完成しヒストグラムを作成しなさい。

階級	度数	累積度数	相対度数	累積相対度数
180以上	10			
170-179	25			
169-169	35			
150-159	25			
149以下	5			

問題 3 データ $x = \{6, 7, 8, 8, 10, 3, 9, 5\}$ として、x の平均、中央値、切り落とし平均、最頻値と四分位点を答えなさい。

問題 4 時系列データ $x_t = \{10, 12, 11, 14, 16, 14, 17\}$ とします。x_t の *4* 項移動平均を計算しなさい。

問題 5 以下の A 社銘柄の株価年次データを利用して、この銘柄の *2007* 年から *2010* 年の間の平均収益率を幾何平均で計算しなさい。

年	2006	2007	2008	2009	2010
株価	2395	2765	2342	2873	3421

問題 6 一定の母集団から無作為に *5* 個のデータをとりだしたところ $\sum_{i=1}^{5}\left(X_i - \bar{X}\right)^2 = 100$ となった。記述統計の分散は平均二乗偏差だからデータの分散は $\boxed{}$ となるが、母集団の分散を求めるには $\hat{\sigma^2} = S^2 = \boxed{}$ を用いるほうがよい。（注意：本資料において以下からデータの分散はすべて不偏の標本分散 S^2 を指す。すなわち $S^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(x - \bar{x})^2$。）

問題 7 *3* 個のデータ

1， 2， 6

の平均値は *3*、標本分散は $\boxed{}$、標準偏差は $\boxed{}$ である。

問題 8 統計学の試験の答案から *6* 人分を抜き取り得点を集計したところ

$$\sum_{i=1}^{6} X_i = 360 \qquad \sum_{i=1}^{6}\left(X_i - \bar{X}\right)^2 = 2400$$

となった。この結果から、平均値と分散はそれぞれ $\bar{X}=$ $\boxed{}$、$S^2=$ $\boxed{}$ と求められる。但し、この *6* 人はサンプル であるので $\hat{\sigma}^2=$ $\boxed{}$ を散らばりの尺度 として採用しておくと母集団の分散を推定する際に不偏性をもつ。

問題 9 *5* 個のデータ

2，　4，　6，　8，　10

の平均値 $\bar{X}$ と標準偏差 S を求めると、$\bar{X}=$ $\boxed{}$、$S=$ $\boxed{}$ になる。ここでデータ全体に *14* を加えると、平均値 は $\boxed{}$、標準偏差は $\boxed{}$ となる。

問題 10 以下の空欄に適当な数値ないし式を記入しなさい。

1. 最近 *1* 週間の読書時間を *5* 名の学生にヒアリングしたところ

3，1，5，5，6（時間）

というデータが得られた。データ全体から $\bar{X}=$ $\boxed{}$、$S=$ $\boxed{}$ になる。

2. 設問 *(1)* の結果を見て「これでは読書時間が足らない」という理由で 全員に *2* 倍の時間の読書を求めることにした。その結果、平均と標準偏差はそれぞれ $\boxed{}$、$\boxed{}$ になる。

問題 11 x の標本分散が c であるとする。$y=2x$ として、y の標本分散を c で表現してください。

問題 12 x の標本分散が 4、平均が 6 とする。x の 4 シグマ区間を計算しなさい。少なくともデータの何パーセントがこの区間に入るかを答えなさい。

問題 13 前の問題の x のデータを標準化しなさい。

問題 14 標準化されたデータの平均が 0、分散が 1 となることを証明しなさい。

問題 15 値が *0*、*1* のみであるデータが、下のような分布表に整理されている。

値	割合
0	*0.6*
1	*0.4*

このデータの平均値は $\boxed{}$、分散は $\boxed{}$、標準偏差は $\boxed{}$ に なる。

問題 16 $x=\{2,4,6\}$、$y=\{6,4,1\}$ とする。x と y の共分散と相関係数を計算しなさい。散布図を書きなさい。

問題 17 テキスト「統計学入門」森棟公夫著 *43* ページ表 *1.17* に基づいて答えなさい。*a.* 就学率が *50%-60%* の国の割合を答えなさい。*b.* 出生率が *20%-30%* の国で就学率が *60%-70%* の国の割合を答えなさい。*c.* 出生率が *20%-30%* の国で就学率が *40%-70%* の国の割合を答えなさい。*d.* 出生率が *20%-40%* の国の割合を答えなさい。

確率分布

問題 18 3 回コインを投げて、$A = \{1$ 回目は表、2 回目は裏、3 回目が裏$\}$ の確率を計算しなさい。

問題 19 2 つのサイコロを振って、出た目の和が 8 以上になる確率を答えなさい。

問題 20 サイコロを 3 回振って、合計が偶数になる順列の数を答えなさい。

問題 21 *10* 円玉を *3* 回投げて、その *10* 円玉の裏表を調べる試行をする。この試行の根元事象と標本空間を書け。

問題 22 {○●△▲☆□ ×} の中から、重複を認めないで *4* 個取り出した時、○が含まれている確率を求めよ。

問題 23 *10* 円玉を *4* 回投げて、その *10* 円玉の裏表を調べる試行をする。この試行の標本空間に含まれる事象の数を求めよ。

問題 24 2つの事象 A, B と条件付確率 $P(B|A), P(B|A^c)$ が与えられているとき、$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)P(A^c)}$ であることを証明せよ。

問題 25 ある国の国民を調査したところ、大学および大学院卒の人が *40* %、それ以 外の学歴の人が *60* %いた。また、大学および大学院卒の人のうち *3* %の人が失業していて、 それ以外の学歴の人のうち 6 %の人が失業していることも分かった。このとき、失業者を調査したとき、その人が大学および大学院卒である確率を求めよ。

問題 26 F_X を一回投げた場合のサイコロの出る目の分布関数とします。F_X のグラフを書きなさい。

問題 27 コインを 2 回を投げる実験を考える。表が出たら 0 裏が出たら 1 とする。2 回を投げた結果の合計の期待値を計算しなさい。

問題 28 X を一回だけのコイン投げの結果（表が出たら 0 裏が出たら 1)、Y をサイコロ一回だけ振った結果（$1, 2, \cdots, 6$）とする。$E(X+Y)$ を計算しなさい。X と Y の実験は互いに独立（無関係に行われた）とする。$E(X \times Y)$ を計算しなさい。

問題 29 10 の目 $(1, 2, \cdots, 10)$ があるサイコロを考える。サイコロの出る目の結果 X の期待値と分散を計算しなさい。

問題 30 成功確率 $p = 0.2$ のベルヌーイ分布の密度関数と分布関数を描きなさい。

問題 31 以下の X の確率分布関数のグラフを描きなさい。

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.2 & 0 \leq x < 5 \\ 0.6 & 5 \leq x < 10 \\ 1 & x \geq 10 \end{cases}$$

問題 32 *3* 枚の硬貨を投げたとき、表が出る硬貨の枚数に興味を持つとする。この実験の確率分布表を作成し、確率関数を図示せよ。

問題 33 離散確率変数 X の確率関数を $p(x_i)$、期待値を μ とする。この確率変数の分散は

$$V(X) = \sum_{i=1}^{n} x_i^2 p(x_i) - \mu^2$$

から計算できることを示せ。

問題 34 確率変数 X は以下のような分布に従う。

値	確率
0	*0.25*
1	*0.5*
2	*0.25*

X の平均値と分散を求めると、$E[X] = \boxed{}$、$V[X] = \boxed{}$となる。

問題 35 歪みのないコインを 5 回投げて 3 回表になる確率を二項分布の分布関数を利用して計算しなさい。

問題 36 ある大学では学生の *3/4* が前期試験の合格者で、残りの *1/4* が他の試験の合格者であるとする。*10* 人の学生を選び出したとき、その中に含まれる前期試験の合格者の数を確率変数 X としたとき、確率関数を図示せよ。また $P(\{X \leq 7\})$ を求めよ。

問題 37 ある南極基地に 1 日平均 1 枚の隕石が落ちるとする。1 週間に 3 枚以上落ちない確率をしめしなさい。

問題 38 連続確率変数 X の確率密度関数を $f(x)$、期待値を μ とする。この確率変数の分散は

$$V(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2$$

から計算できることを示せ。

問題 39 X が自由度 6 の t 分布に従うとする。t 分布表を利用して $P(X \geq 1.44)$ を求めなさい。

問題 40 統計学の試験を行ったところ得点分布には正規分布 $N(60, 12^2)$ が概ね当てはまっていた。受験生は *400* 人だった。以下の設問に答えなさい。得点が *42* 点以下の者は全体の何*%* 程度と考えられますか。 A 君の得点は *72* 点だった。受験生全体の中で A 君の得点順位は何位くらいと推測されますか。

問題 41 日本人全体の身長分布は正規分布 $N(165, 225)$ で表せることがわかっているとする（単位：*cm*)。以下の空欄に適当な数値を記入しなさい。

1. 日本人の身長の 1 シグマ区間は$\boxed{}$*cm* から$\boxed{}$*cm* までの範囲となる。
2. 身長が *195cm* 以上の人は全体の$\boxed{}$*%* 程度いるはずである。
3. *165cm* から *180cm* までの範囲に属する人は全体の$\boxed{}$*%* 程度いるはずである。

問題 42 統計学の試験を採点したところ得点分布は正規分布 $N(65, 15^2)$ が当てはまることがわかった。以下の空欄に適当な数値を記入しなさい。

1. *70* 点の答案を標準値に直すと$\boxed{}$になる。
2. *80* 点以上の者は全体の$\boxed{}$%程度いるはずである。
3. この得点分布における 1 シグマ区間は$\boxed{}$点から$\boxed{}$までの範囲である。*1* シグマ区間に含まれる受験者は全体の$\boxed{}$%程度を占めるはずである。

問題 43 変数 X について平均と分散がそれぞれ $E[X]=10$、$V[X]=100$ であることがわ かっている。このとき、$Y=2X+1$ のように変数 Y を定義すると、 $E[Y]=\boxed{}$、$V[Y]=\boxed{}$となる。また、この 二つの結果から $E\left[Y^2\right]=\boxed{}$となることもわかる。

問題 44 変数 X について $E[X]=0$、$SD[X]=1$ が分かっている。変数 Y を $Y=-2X$ と定 義すれば、$E[Y]=\boxed{}$、$SD[Y]=\boxed{}$になる。

問題 45 互いに独立（＝無関係）な変数 X、Y について、$E[X]=2,\ V[X]=4$、および $E[Y]=-2,\ V[Y]=9$ がわかっている。このとき、$E[X+Y]=\boxed{}$になる。また、独立ということから $V[X+Y]=\boxed{}$である。

問題 46 統計学の試験を受験した学生全体について得点分布を調べると $N\left(60,15^2\right)$ が 当てはまっていた。ここで得点を標準値 Z に変換すると分布の中の 位置づけがよくわかる。たとえば、*80* 点を標準化すると $\boxed{}$になる。偏差値 $S.S$ というのは $S.S=10\,Z+50$ のよう に定義される値である。だから $S.S$ の分布の平均と標準偏差はそれ ぞれ$\boxed{}$、$\boxed{}$となる。

問題 47 確率変数 Z は標準正規分布にしたがっているなら、$P(Z>1.82)=0.\boxed{}$である（注、数値表に記載されている値の 最初の *4* 桁を記入すること）。

問題 48 平均が *130*、標準偏差が *20* である正規分布において、値が *110* から *170* までにおさまる割合は$\boxed{}$*%* である。

問題 49 確率変数 X について $E[X]=2$、$V[X]=9$ が分かっている。 このとき、$Y=2X-3$ で変数 Y を定義すると $E[Y]=\boxed{}$、$V[Y]=\boxed{}$となる。

問題 50 平均と分散がそれぞれ *60*、*1* である母集団から無作為に *10* 個のデータを取りだす とすれば、$E\left[\bar{X}\right]=\boxed{}$、 $SD\left[\bar{X}\right]=\boxed{}$となることがわかる。

問題 51 平均と分散がそれぞれ *170*、*100* である母集団からランダムに *20* 個 のデータをとって標本平均 $\bar{X}$ を求める。このとき $\bar{X}$ の分布につい て、$E\left[\bar{X}\right]=\boxed{}$、 $SD\left[\bar{X}\right]=\boxed{}$になることがわかる。

問題 52 互いに独立な確率変数 X と Y について $E[X]=0,\ V[X]=1$ および $E[Y]=10,\ V[Y]=5$ が分かっている。 このとき $E\left[\frac{X+Y}{2}\right]=\boxed{}$となり、 $V\left[\frac{X+Y}{2}\right]=\boxed{}$となる。同じようにすれば $V\left[\frac{X-Y}{2}\right]=\boxed{}$になることもわかる。

問題 53 確率変数 X について $E[X]=4,\ V[X]=4$ であることが分かっている。このとき $E[X-4]=\boxed{}$となり $V[X-4]=\boxed{}$となる。また X^2 の平均値 $E\left[X^2\right]$ は$\boxed{}$になることもわかる。

問題 54 互いに独立な確率変数 $X,\ Y$ について $E[X]=4$、$V[X]=8$ および $E[Y]=-4$、$V[Y]=8$ が分かっている。このとき $E[X+Y]=\boxed{}$、$V[X+Y]=\boxed{}$となる。したがって、$E\left[\left(\frac{X+Y}{4}\right)^2\right]=\boxed{}$になる。

問題 55 確率変数 X は正規分布 $N(2,16)$ にしたがっている。このとき確率 $P(X \geq 3) = \boxed{}$ になる。確率変数 Y を $Y = -2\,X$ と定義 すれば、その平均と分散はそれぞれ $E[Y] = \boxed{}$、 $V[Y] = \boxed{}$ になる。$E\left[Y^2\right] = \boxed{}$ であ ることもわかる。変数 Y の標準化とは

$$Z = \frac{Y - \boxed{}}{\boxed{}}$$

のように変換することをいう。標準化を行うと常に平均は $\boxed{}$、標 準偏差は $\boxed{}$ となる。

問題 56 $V[X] = E\left[X^2\right] - \{E[X]\}^2$ であることを示しなさい。

問題 57 以下の空欄に適当な数値、式、語句を記入しなさい。 Z_1, Z_2, Z_3, Z_4, Z_5 はいずれも標準正規分布にしたがう互いに独立な確率変数 である。$X = Z_1 + Z_3 + Z_5$ で定義される確率変数 X の分布は平均が $\boxed{}$、分散が $\boxed{}$ の $\boxed{}$ 分布であ る。したがって、$\dfrac{(Z_1 + Z_3 + Z_5)^2}{3}$ の値が示す分布は $\boxed{}$ 分布となり、平均値は $\boxed{}$ となる。また、この分布の分散に着目すれば、 $V\left[(Z_1 + Z_3 + Z_5)^2\right] = \boxed{}$ であること がわかる。

問題 58 互いに独立な確率変数 X と Y について $E[X] = -1$、$V[X] = 4$、$E[Y] = 1$、$V[Y] = 4$ であることがわかっている。この 時、$E[X+Y] = \boxed{}$ となり、また X と Y は独立だか ら $V[X+Y] = \boxed{}$ となる。このことは $\dfrac{(X+Y)^2}{8}$ の値が示す分布が $\boxed{}$ 分布であり、 その平均が *1* であることと関係する。この分布の分散の性質に注意 すれば、$V\left[(X+Y)^2\right] = \boxed{}$ であることがわかる。

問題 59 確率変数 X は平均が *2*、分散が *9* の正規分布にしたがっている。この時、 $Y = \dfrac{X-2}{3}$ で定義される変数 Y の分布は、 平均が $\boxed{}$、分散が $\boxed{}$ の $\boxed{}$ 分布となる。したがって、Y^2 の分布は $\boxed{}$ 分布とな り、$E\left[Y^2\right] = \boxed{}$、$V\left[Y^2\right] = \boxed{}$ とな る。

問題 60 日本人全体の身長分布は正規分布 $N(165, 225)$ で表せることがわかっていると する（単位：*cm*）。以下の空欄に適当な数値を記入しなさい。

問題 61 日本人の身長の 1 シグマ区間は $\boxed{}$ *cm* から $\boxed{}$ *cm* までの範囲となる。 身長が *195cm* 以上の人は全体の $\boxed{}$ % 程度いるはずである。 *165cm* から *180cm* までの範囲に属する人は全体の $\boxed{}$ % 程度いるはずである。

問題 62 確率変数 X は正規分布 $N\left(130, 20^2\right)$ に従っている。値 $X = 145$ を標準化 すると $Z = \boxed{}$ になるので、確率 $P(X \leq 145)$ は $\boxed{}$ である。同様に、$P(140 \leq X < 150) = \boxed{}$ である。

問題 63 身長の母集団分布は正規分布 $N(168, 100)$ であることがわかっている。母集団から 無作為に *10* 人を抽出し身長の測定値 $X_1, \ldots, X_{10}$ を得る。データから標本分散 S^2 を求めるとして設問に答えなさい。

1. S^2 の標本分布の平均値を回答しなさい。

2. S^2 の標本分布の分散を回答しなさい。

3. $P\left(S^2 \geq 100\right)$ は 0.5 に達しないことを示しなさい。S^2 の分布に関してこのことが意味する内容を簡単に述べなさい。

問題 64 変数 X の分布は $N(0,1)$、変数 Y の分布は $N(1,2)$ であり、X と Y は互いに独立とする。このとき、$E[X+Y]=\boxed{}$、 $V[X+Y]=\boxed{}$ となる。変数 W を $W=X^2+\left(\frac{Y-1}{\sqrt{2}}\right)^2$ と定義すると、W の 分布は $\boxed{}$ であることが分かるので、W の平均と分散は それぞれ $E[W]=\boxed{}$、$V[W]=\boxed{}$ となる。

問題 65 以下の空欄に適当な語句、数値、記号を入れなさい。確率変数 X は正規分布 $N(-1,1)$、Y は $N(1,1)$ にしたがい、互いに独立 とする。このとき $X+Y$ の平均と分散は、それぞれ $E[X+Y]=\boxed{}$、$V[X+Y]=\boxed{}$ となる。ま た、$X+1$ の分布は $\boxed{}$ となるので、それを二乗した $(X+1)^2$ は $\boxed{}$ 分布にしたがうことになる。同様に $W=(X+1)^2+(Y-1)^2$ で変数 W を定義すれば、W は $\boxed{}$ 分布にしたがうことが分かるので、 $E[W]=\boxed{}$、$V[W]=\boxed{}$ となる。

問題 66 データ分析では分布という視点から問題を考えるが、データの分布で ある $\boxed{}$ 分布と理論的な分布である $\boxed{}$ 分布を区別することが大事である。理論的な分布をとりあげるときには値を X とし て平均を $\boxed{}$、分散を $\boxed{}$ という記号で表す。

問題 67 確率変数 X は正規分布 $N(1,1)$ にしたがっている。この時、確 率 $P(X \geq 1)=\boxed{}$ であり、また $P(0 \leq X \leq 1)=\boxed{}$ となる。

問題 68 正常なコインを投げて表が出れば *1*、裏が出れば *0* とする。いま実際にコインを投 げてみると表が *13* 回出た。以下の設問に解答しなさい。

1. データの平均（＝標本平均）とデータの分散（＝標本分散）を求めな さい。

2. データが従っている確率分布の平均と分散を求めなさい。

問題 69 統計学の試験を受けた学生全体の得点は正規分布 $N(55,225)$ であることがわかっ ている。以下の設問に解答しなさい。

1. 得点分布を図に描き、図の中に平均と標準偏差を書き入れなさい。更 にこの分布の「*1* シグマ区間」が分かるように薄く斜線を入れなさい。

2. *40* 点以下の学生は全体の何% 程度いるはずですか。

3. 評価が良の学生は全体の何% いますか。但し、得点が *70* 点以上 *80* 点以 下が良になります。

4. 学生全体に一律に *10* 点のゲタをはかせ得点調整をすることにした。そ の結果、成績が不可の学生の割合は何% から何% に減りますか。但 し、成績が不可になるのは得点が *60* 点以下の場合です。

問題 70 血圧検査を受けた受診者全体について血圧値 (X) の分布を見ると正規分布 $N(130, 20^2)$ が当てはまっている。以下の設問に解答しなさい。

1. 血圧分布を図に描き、図の中に平均と標準偏差を書き入れなさい。更 にこの分布の「*1* シグマ区間」が分かるように薄く斜線を入れなさい。
2. 血圧値が *100* 以下である受診者は全体の何%いるはずですか。
3. この分布の *2* シグマ区間とそこに含まれる受診者の割合を答えなさい。
4. 血圧値が *120* から *140* までの受診者は全体の何%いるはずですか。
5. コップ *1* 杯のビールを飲んで *10* 分後に測定すると血圧が *10%* 上昇する とします。このとき設問 *(4)* で求めた $P(120 \leq X \leq 140)$ は どのように変わりますか。

問題 71 正常なコインを *2* 個用意して投げるとする。コイン A が表なら $X = 1$、 裏なら $X = 0$ とする。同様にコイン B が表なら $Y = 1$、裏なら $Y = 0$ とする。$X + Y$ の 確率分布を考えると $E[X + Y] = \boxed{}$、$V[X + Y] = \boxed{}$である。ということは $E\left[(X + Y)^2\right]$ もわかり$\boxed{}$になる。

問題 72 確率変数 X は正規分布 $N(0, 9)$ にしたがっている。変数 Y を $Y = -2X + 1$ のように定義すると $E[Y] = \boxed{}$、$V[Y] = \boxed{}$になる。また、$E\left[X^2\right] = \boxed{}$、$E\left[Y^2\right] = \boxed{}$にな ることも容易に確かめられる。

問題 73 標準正規分布 $N(0, 1)$ から *5* 個の無作為標本 $Z_1, Z_2, \ldots, Z_5$ を抽出する。 このデータの標本平均 $\bar{Z}$ がしたがう分布は$\boxed{}$で あるので $P\left(\bar{Z} \geq 0.3\right) = \boxed{}$であること が数値表からわかる。また $W = \sum_{i=1}^{5} Z_i^2$ の値も分布するが、 この分布の形は$\boxed{}$と呼ばれるものである。変数 W の平 均と分散は $E[W] = \boxed{}$、 $V[W] = \boxed{}$となることがわかっている。

問題 74 ある自治体の失業者に関する分析結果から以下のことが分かった。失業者が一年以内で再就職できる確率は 30%、職業訓練を受けられる確率が 36%、一年以内で再就職できる者が職業訓練を受ける確率が 60% となる。失業者になった人が職業訓練を受けて一年以内で再就職できる確率を求めなさい。

標本分布

問題 75 正規分布 $N(120, 15^2)$ を母集団として *1* 個のデータを無作為に取り出すと値 130 が得られた。この値を標準値 Z に直すと$\boxed{}$になる。ま た同じ母集団から無作為に *10* 個のデータを抽出すると $\bar{X}_{10} = 130$ になった。この値を標準化すると$\boxed{}$ になる。このように母集団の分布とデータの結果がしたがう標本分布 を区別することが重要である。この標本平均 $\bar{X}_{10}$ の 標本分布を考えると $P\left(\bar{X}_{10} \geq 125\right) = \boxed{}$であることが数値表からわ かる。

問題 76 ある航空機の定員は *256* 人である。この航空機が満席時の旅客総体重は最も重い 時で何キロ程度と予想しておけば十分ですか。但し、この航空機を 利用する人たちの体重分布は $N(50, 16^2)$ とする。（注）十分だ と考える確率的な根拠も付け加えること。

問題 77 ある銘柄の株価の毎日の変動は、平均がゼロ、標準偏差が *100* 円の分布に従ってお り、毎日の騰落は互いに独立で、規則性はないことが知られている。 今日の終値は *7000* 円である。明日以降 *100* 営業日が経過した時点で、こ の株価はいくらになるかを予想したい。設問に答えなさい。

1. 今日から *100* 日間の毎日の価格変化を合計した値を確率的に予想したい。*100* 日間の 変化の合計を X として、それに当てはまる確率分布を図に描き、そ の期待値と標準偏差を図の適切な箇所に記しなさい。

2. *100* 営業日後の終値が *4000* 円を下回っている可能性は、十分あると考慮するべきか、確率的にはほとんどないかを答えなさい。理由も書 くこと。

問題 78 あるスーパーのレジで、1人の客に対して清算を始めて終わるまでの時間には 正規分布 $N\left(180, 30^2\right)$ が当てはまっている。(単位：秒)。 以下の設問に答えなさい。

1. レジで清算する時間が *4* 分を超える客は全体の何パーセント位いるでしょう。(*5* 点)

2. *9* 人の客をサンプルにとってレジで清算する時間を計測する。*9* 人の平均時間 $\bar{X}$ に当てはまる確率分布を図に描き、図の中の適切な 場所に期待値 $E\left[\bar{X}\right]$ と標準偏差 $SD\left[\bar{X}\right]$ の値を書き 入れなさい。

3. 自分の番が回ってくるまでに *9* 人の客がいる。最長でどの程度待てば自分の番が回ってくるでしょうか。

問題 79 母集団分布が $N\left(115, 20^2\right)$ だとする。*16* 個のサンプルを無作為にとって平均値 $\bar{X}$ を求めるとする。以下の設問に答えなさい。

1. 平均値 $\bar{X}$ が従う確率分布を下に描き、図の適当な場所に $E\left[\bar{X}\right]$ と $SD\left[\bar{X}\right]$ の値を書きなさい。

2. 確率 $P\left(\bar{X} \geq 125\right)$ を求めなさい。(有効数字 4 けた)

3. 確率 $P\left(110 \leq \bar{X} < 120\right)$ を求めなさい。(有効数字 4 けた)

問題 80 *2005* 年度における東証一部上場企業全体の増益率は正規分布 $N\left(15, 5^2\right)$ のように分布しているとする（単位：%)。いまラン ダムに *10* 社をサンプルにとり増益率を確かめる。以下の設 問に答えなさい。

問題 81 サンプルの平均値 $\bar{X}$ の標本分布を図に描き、その平均と標準偏差を図の中に書き入れなさい。(小数の場合は第 *2* 位まで) 設問 *(1)* の分布において $\bar{X} = 17$ を標準値にしなさい（小数第 *2* 位 まで)。平均値 $\bar{X}$ の値が *17%*を超える確率を求めなさい（小数第 *4* 位まで)。

問題 82 母集団 $N(120, 100)$ から無作為に *10* 個のデータ $X_1, X_2, \ldots, X_{10}$ をと りだして標本平均 $\bar{X}$ を求める。このとき確率 $P\left(\bar{X} \geq 125\right)$ は $\boxed{}$ である。ここで同じ 母集団からとりだすデータ数を *10* 個から *20* 個に増やして標本平均を求めるとすれば、$P\left(\bar{X} \geq 125\right) = \boxed{}$ になる。デー タ数を増やすと標本平均は分布の中心に集まってくることがわか る。

問題 83 ある国民の身長分布は平均が *168cm*、標準偏差が *10cm* の正規分布にしたがってい ることが知られている。以下の設問に解答しなさい。

1. この国民全体を母集団として無作為に抽出する *1* 名の身長を X *(cm)* とする。確率 $P(158 \leq X \leq 178)$ はいくらですか。

2. 設問 *(1)* でいう確率変数 X から $Z = (X - 168)/10$ のように確率変数 Z を定義する。 このとき確率 $P(Z \leq -1)$ はいくらになりますか。

3. この母集団から無作為に *5* 人の標本 $X_1, X_2, \cdots, X_5$ を抽出したうえで標本 平均 $\bar{X}$ を求めるとする。

(a) $\bar{X}$ がしたがう標本分布の形を図に描き、その平均と標準偏差 の値を図の中に書き加えなさい。

(b) 標本平均 $\bar{X}$ の値が *170cm* から *175cm* までの範囲に入る確率はど の程度あるかを答えなさい。

問題 84 統計学の試験の解答状況を把握するため無作為に *10* 名の受験者を抽出し平均 $\bar{X}$ を求める。実は受験者全体の得点分布は正規分布 $N\left(60, 15^2\right)$ になっている。*10* 名の標本平均 $\bar{X}$ はどんな分 布に従いますか。分布図を描いたうえ、$\bar{X}$ の平均と標準偏 差を図の中に書き加えなさい。

問題 85 データ数がある程度大きければ $\bar{X}$ の標本分布は正規分布で近似できる。 社会全体の視聴率が *30%* である番組について、「みた（*1*)」か「みなかった （*0*)」かというアンケート調査を *100* 人を対象に行うとき、「みた」 と回答する人の割合は、平均が$\boxed{}$、標準偏差が $\boxed{}$ の正規分布に従う。したがって、アンケート結果 を *3* シグマ区間で憶測すると高々$\boxed{}$までの値である。(回答するとき割合は小数で表されるものとしなさい)

問題 86 統計学の試験の得点分布は正規分布 $N\left(60, 15^2\right)$ である。無作為 に *6* 人を抽出して得点をしらべ標本分散 S^2 を計算する。設問に答 えなさい。

1. 統計量 S^2 の標本分布の平均を回答しなさい。
2. 統計量 S^2 の標本分布の分散を回答しなさい。
3. S^2 とは別に「分散＝偏差二乗の平均」に忠実に計算するため $S_*^2 = \frac{1}{6}\sum_{i=1}^{6}(X_i - \bar{X})^2$ で定義される値 S_*^2 を利用することにした。S_*^2 の標本分布の平均 を求めなさい。

問題 87 確率変数 X は正規分布 $N(168, 100)$ にしたがっている。この分 布から無作為に *10* 個のデータ $X_1, X_2, \ldots, X_{10}$ を抽出し、標本平均 $\bar{X}$ を求めるとすれば、$\bar{X}$ の標本分布は $N\left(\boxed{}, \boxed{}\right)$ となることがわかる。つまり、$E\left[\bar{X}\right] = \boxed{}$ となるが、$\bar{X}$ の持つこの性質を$\boxed{}$という。

問題 88 *5* 個のデータ *1*、*2*、*3*、*4*、*5* の分散を計算する場合、分散＝偏差二 乗値の平均という「分散」の定義に基づいて求めると$\boxed{}$となるが、 集団全体の真の分散 σ^2 を不偏推定する立場にたてば$\boxed{}$とな る。

問題 89 正規分布 $N(1, 1)$ から無作為に *5* 個のデータを抽出し、 $\sum_{i=1}^{5}(X_i - 1)^2$ を計算する場合、この式の値はデータ にもとづく確率変数となり、その分布は$\boxed{}$と呼ばれている。更に、 この分布の平均値は$\boxed{}$である。

問題 90 ある生産ラインで生産される蛍光灯の寿命は一定の形の分布をして おり、平均寿命 μ は *2000* 時間、標準偏差 σ は *150* 時間である。この生産 ラインから無作為に *100* 個の製品を抜き取り、寿命の検査をすることにした。以下 の設問に解答しなさい。

1. 抜き取った *100* 個の平均寿命を $\bar{X}$ とするとき、$\bar{X}$ はどのような分布にしたがいますか。分布の形を図に描き、その平均値と標準偏差を求めなさい。
2. 確率 $P\left(1990 \leq \bar{X} \leq 2010\right)$ を求めなさい。

問題 91 標本平均の分布の知識は統計学の基礎である。たとえばサイコロを振って出る目の数を X とすれば $E[X] = 3.5$、$V[X] = 2.92$ になる。いまサイコロ を *50* 回振って出た目の数の平均値を計算するとしよう。この平均値の 1 シグマ区間は$\boxed{}$から$\boxed{}$であると言ってよ い。この 1 シグマ区間に可能性全体の$\boxed{}$*%* が含まれる。

問題 92 正しいサイコロを振って出る目の数を X とすると $E[X] = 3.5$、$V[X] = 2.92$ になる。いま正しいサイコロを *100* 回振って出る 目の数の平均値 $\bar{X}$ を求めると、$\bar{X}$ の値には正規分 布 $N\left(\boxed{}, \boxed{}\right)$ が概ね当て はまる。

問題 93 正常なサイコロを振って出る目の数を X とすると $E[X] = 3.5$、$V[X] = 2.92$ と なることがわかる。いまこのサイコロを *100* 回振って、目の数 の平均値 $\bar{X}$ を求めるとする。以下の設問に答えなさい。

1. 平均値 $\bar{X}$ はどのような形の標本分布に従うか。下に分布図を描き、 図の中に分布の中心と散らばりの度合いを書き入れなさい。
2. この平均値が示す値として最大値と考えられる値は、確率的にどの程度の値か。 そう思われる理由も書き添えて回答してください。

問題 94 母集団 $N\left(120, 15^2\right)$ から無作為に *20* 個のデータ $X_1, X_2, \ldots, X_{20}$ をとり

$$\frac{1}{19}\sum_{i=1}^{20}\left(X_i - \bar{X}\right)^2$$

を計算する。この式の値を W とするとき W の標本分布における $E[W]$ と $V[W]$ を求めなさい。

問題 95 サンプルサイズ $n = 16$ のデータの標本平均 $\bar{X} = 9$、標本偏差 $S = 3$ とする。95% の信頼区間を求めなさい。

問題 96 サンプルサイズ $n = 100$ のデータの標本平均 $\bar{X} = 9$、標本偏差 $S = 3$ とする。99% の信頼区間を求めなさい。

問題 97 正規母集団 $N\,(50, 169)$ から無作為抽出された、13 個の標本の標本平均 $\bar{X}$ について $P\left(\left\{\bar{X} \geq 40\right\}\right)$ と $P\left(\left\{45 \leq \bar{X} \leq 55\right\}\right)$ を求めよ。

統計的推測：推定

問題 98 標本平均が期待値の不偏推定量であることを証明しなさい。

問題 99 ある菓子メーカーが製造している菓子から *9* 袋のサンプルをとり重さを測ったところ、以下のようになった（単位：グラム）。

51	*49*	*50*	*54*	*51*
50	*48*	*52*	*52*	

データを要約すると、$\bar{X} = 50.8$、$\hat{\sigma} = 1.79$ だった。 菓子袋全体の平均値を μ として、μ の値を推定しなさい。信 頼係数は *95%* とする。

問題 100 ある人がレストランを開店しようとしている。予定地の近くで食事をした人 *25* 人 に支払った金額をきいたところ、結果は $\bar{X} = 1200$、不偏 分散 $\hat{\sigma^2} = 300^2$ となった *(*単位：円)。客層となる 母集団全体では平均的にいくら位の支払いを考えているかを推定し なさい。信頼係数は *95%* とする。

問題 101 ある弁当屋のライスを（ランダムに）*10* 個買ってきて重さを測ったところ

$$\sum_{i=1}^{10} X_i = 2490 \qquad \sum_{i=1}^{10}\left(X_i - \bar{X}\right)^2 = 170$$

であった（X_i は個々の重さを示す）。この店で売られているライス全体とし ては、どのくらいの平均重量になっていると推定されますか。（注） 点推定値、標準誤差の大きさ、信頼係数を示したうえで解答するこ と。

問題 102 ある高校からランダムに *64* 名をとって *100* メートル走の記録を確 認したところ、*64* 名の平均と分散がそれぞれ $\bar{X} = 13.2$、 $\hat{\sigma^2} = 1.44$ だった。学年全体の平均タイムを μ とお くとき、μ の値を信頼係数 *90%*で推定しなさい。

問題 103 正規母集団から無作為にとった 7 個のデータから平均と分散を計算すると

$$\bar{X} = 170.5 \quad \hat{\sigma^2} = 25.0$$

だった。 このデータが集められた母集団の平均 μ がいくらであるかを信頼係数 *95%* で区間推定する。結果は

$$0.95 = P\left(\boxed{} \le \mu \le \boxed{}\right)$$

となる。

問題 104 無作為に *5* 個のデータをとると、たまたま

1, 2, 3, 4, 5

というデータが得られた。以下の設問に回答しなさい。

問題 105 標本平均は、当然、$\bar{X} = 3$ になるが、データから分散を求めるには二通りある。分散の定義どおりに平均二乗偏差を求めると $S^2 = \boxed{}$ になる。他方、不偏分散を求めるなら $\boxed{} = \boxed{}$ という値になる。 今回のデータから母集団の平均値 μ を推定する。信頼係数は *95%* とすると、以下のような計算をすればよい。自由度が $\boxed{}$ の T 分布を利用すればよいので、

$$\begin{aligned}
&0.95 \\
=&P\left(-\boxed{} \le T \le \boxed{}\right) \\
=&P\left(\boxed{} - \boxed{} \times \sqrt{\frac{\boxed{}}{\boxed{}}} \le \mu \le \boxed{} + \boxed{} \times \sqrt{\frac{\boxed{}}{\boxed{}}}\right)
\end{aligned}$$

問題 106 全国学力試験を行った。数学の得点分布を見るために無作為に *10* 名の得点を抽 出すると平均点が $\bar{X} = 55$ となった。以前、全国学力試験を実 施していた期間の得点分布から数学の得点の標準偏差は *20* 点前後と ほぼ一定であったことがわかっている。今回も母集団の標準偏差は *20* 点としよう。さて、全国の受験者全体の平均点はいくら程度と推定でき るか。信頼係数を明示のうえ信頼区間を示しなさい。その信頼区間を踏まえ「全国の平均点が *70* 点に達する」と考えておいていいかど うかにも触れること。

問題 107 『母分散 σ^2 をできるだけ正確に推定したいとき無作為データから計算し た S^2 を用いるのは不適切である』という指摘の意味を丁寧に説明 しなさい。説明の中でデータ数を n 個と置きなさい。

問題 108 データに基づいて推定や検定を行うとき、無作為に抽出されるデータのことを $\boxed{}$ と言い、データ抽出の対象となる範囲のことを $\boxed{}$ と呼んでいる。また、$\bar{X}$ や S^2 のようなデー タの関数を一般に $\boxed{}$ と言っている。$\boxed{}$ の標 本分布は推定や検定を行う際の理論的な基礎になる。

問題 109 ロケットの打ち上げ実験を *30* 回行ったところ *24* 回打ち上げに成功し、 *6* 回失敗した。技術上定まっている成功率を未知の値 p として、以 下の設問に答えなさい。

1. 30 回の実験で X 回成功したときの成功率を $\hat{p}$ と表すことに する。成功率 $\hat{p}$ の標本分布を図に描き、近似できる分布型の 名称と平均、標準偏差の値を図に書き加えなさい。

2. $\hat{p}$ の標準偏差を概数で求めなさい。

3. 未知の値 p を信頼係数 *95%* で区間推定しなさい。

問題 110 ある企業の従業員を母集団として血圧分布の状況を調べたい。無作 為に選んだ *10* 人の血圧を測定し、標本平均をとると $\bar{X} = 124$ となった。血圧の標準偏差は過去の経験から $\sigma = 10$ であることが わかっているので、母平均 μ について推定を行う。設問に答えな さい。

1. *10* 人の平均 $\bar{X}$ の標本分布はどんな形となるか、それを図に描き、分布型の名称と平均、標準偏差を図に書き加えなさい。

2. 信頼係数を *95%* として、母平均 μ を区間推定しなさい。

3. 信頼係数を *95%* とし、± 2 の精度で母平均 μ の値を区間推定し たい。最低何人の従業員からデータを得ればよいかを回答しなさい。

問題 111 乾電池を生産している企業が新製品を開発した。無作為に *15* 個の新製品を抜き取 り、所定の方式で耐久時間の測定をしたところ、結果は $\bar{X} = 196.0$ （時間）、$\hat{\sigma^2} = 900.0$ となった。旧型の製品の耐久時 間は平均 180.0（時間）、標準偏差は 25 であることがわかっている。 以下の設問に解答しなさい。

1. 新製品の耐久時間のばらつきは旧型の製品と同じと仮定してよいとす る。このとき新製品の耐久時間は平均いくら程度と思われるか。今回の測定結果に基づいて区間推定をしなさい。区間推定を行うときの信頼 係数も記すこと。また、旧型と比較した場合の新型乾電池の耐久性をどの程度まで評価できるか思うところを簡単に記しなさい。

2. 新製品の平均耐久時間を点推定しなさい。

3. 区間推定と点推定の二つを比較考察し、点推定よりも区間推定のほう が方法として優れていると思われる理由を簡潔に述べなさい。

統計的推測：検定

問題 112 ある新聞に以下のような記事が掲載された。『ある自動車メーカーが発表した新モデル車は、高速で *1* リットル当たり平均 *30* キロ、 標準偏差が *6* キロで走ると説明している。しかし、消費者グループが 同型車を *9* 台、高速で走らせたところ、燃費の平均値は *1* リットル当 たり *27* キロだった。消費者グループはメーカーの説明は正しくない と主張している。』あなたは消費者グループの主張は正しいと考えますか。そう考える理由と併せて 答えなさい。

問題 113 ある工業団地を抱える市では、「雨水の pH 値は *5.5* で心配はない」と 市当局が言っている。そこで、市民団体が *10* 地点で雨水の pH 値を測定したところ、平均が *5.28*、標準偏差が *0.215* になった。次の設問に答えなさい。空欄がある場合 は適当な数を記入し、選択肢が与えられている問題では番号を丸で囲むこと。

1. 市の言い分が正しいとすれば、今回の測定結果では $\boxed{}$ の誤差が出たことになる。

2. サンプルにとった数とデータの散らばりから考えて、平均値には$\boxed{}$の標準誤差があると見られる。

3. 市の言い分が正しいとして、今回のサンプルは（1．高い確率で予想 できる、2．まず予想できない)。以上により、市の言い分は（1．信頼しても よい、2．信頼できない）と判断できる。

4. 上の設問 *(3)* で得た判断が誤っているとすれば、それは（1．第 *1* 種 の誤り、2．第 *2* 種の誤り）に該当する。

問題 114 統計学の試験終了後、ランダムに *8* 人の答案を抜き取り平均値 $\bar{X}$ と不偏 分散値 $\hat{\sigma^2}$ を計算すると、

$$\bar{X} = 55 \quad \hat{\sigma^2} = 121$$

になった。受験者全体の平均点 μ が *60* 点を下回っているとすれば採点を甘めにつけようと考えている。上のサンプルの結果から全体の平均点が *60* 点未 満であると判断してよいかについて以下の設問に答えなさい。

1. 今回判断したいことを統計的検定の帰無仮説と対立仮説に定式化すると

$$H_0 : \boxed{} \quad vs \quad H_1 : \boxed{}$$

   のように書ける。

2. 統計的検定の手順に従って、今回の標本平均から T 値を求めると

$$T_0 = \frac{\boxed{} - \boxed{}}{\sqrt{\boxed{} \Big/ \boxed{}}} = \boxed{}$$

   になる。

3. 有意水準を *5%* として$\boxed{}$側に棄却域を設ける。限界値は数値表から$\boxed{}$となる。したがって、帰無仮説は$\boxed{}$。検定の 結論としては、受験者全体の平均点は$\boxed{}$
   但し、今回の判断には第$\boxed{}$種の過誤をおかしている可能性がある。

問題 115 あるメーカーの乗用車のブレーキは時速 *60Km/h* から急ブレーキをか けた時に *60* メートルで停止するように設計されている。このブレー キについて以下の検定問題をとりあげる。次の設問に答えなさい。 なお、μ は停止距離の母平均。

$$H_0 : \mu = 60 \quad vs \quad H_1 : \mu = 62$$

1. 実際に *8* 回反復して停止実験を行い結果を集計すると $\bar{X} = 61.5$ メートル、不偏分散は $\hat{\sigma^2} = 3$ だった。この結果に基づく と、ブレーキについてどのように判断をすればよいか答えなさい。有意水準は *5%* 。また毎回の停止距離には標準偏差で *2* メートルのばらつ きがあることが分かっている。解答するときは、限界値を示したうえ で、結論を下すこと。

2. 設問 *(1)* で下した結論には、どのような判断ミスの危険がどの程度あるかを 簡潔に述べなさい。

問題 116 ある工場に勤務する *20* 歳台の者から *10* 人を無作為に選び血圧を測定したところ、 平均 $\bar{X}$ が *138*、分散推定値 $\hat{\sigma^2}$ が *95* になった。 その年齢層全体の平均血圧 μ が *135* を超えているとすれば、原因を詳し く調べたい。今回の *10* 人の結果から調査の必要があるかどうかを検 定することにした。

1. 検定問題を帰無仮説と対立仮説に定式化すると

$$\mathrm{H}_0 : \boxed{} \quad vs \quad \mathrm{H}_1 : \boxed{}$$

   になる。

2. 帰無仮説を正しいと仮定した上で、今回の標本平均である $\bar{X} = 138$ を標準化すると

$$T_0 = \frac{\boxed{} - \boxed{}}{\sqrt{\boxed{} \Big/ \boxed{}}} = \boxed{}$$

   となる。

3. 結論としては、帰無仮説を（1．採択する、2．棄却する）。 データから帰無仮説を棄却できないからといって、それは帰無仮説が正しいとい うことの根拠に必ずしもならないことに注意する必要がある。そ れは第$\boxed{}$種の過誤をおかしているかもしれないからである。 検定をするときには、上に指摘した過誤をおかす確率を 1 か ら差し引いた$\boxed{}$を高くすることが大切である。

問題 117 統計的な仮説検定は、形の上では二択問題である。たとえば、ある 工場に勤務する *20* 歳台の従業員全体の平均血圧 μ が *140* を超えていれば、残業 の常習化、食生活の悪化など原因を詳しく調べたい。そのためランダムに *10* 人を 選んで血圧を測定する。統計学ではそれぞれの選択肢を仮説と呼ぶが、上の例では一方の仮説を $\boxed{}$、他方の仮説を$\boxed{}$とすればよい。基本的な考え方としては、前者の仮説を正しいと仮定し、データから求められる 平均値 $\bar{X}$ が限界値を超えれば後者、超えなければ前者をとる。用語としては前者の仮説を$\boxed{}$仮説、後者の仮説を$\boxed{}$仮説という。 前者の仮説を否定する棄却域の大きさとして、通常は *5%* をとるが、確率で示した 棄却域の大きさを$\boxed{}$と呼んでいる。

問題 118 ある高校の *1* 年生から無作為に *9* 名を選び *100* メートル走の記録をとると標本平均 と分散推定値がそれぞれ

$$\bar{X} = 13.5 \qquad \hat{\sigma^2} = 1.44$$

になった。但し、毎年の経験から学年全体の分散は今回も *1.8* だとわかっ ているものとする。以下の設問に解答しなさい。

1. 学年全体の平均タイムはどの程度と思われるか。信頼係数 *95%*で区間推定しなさ い。

2. 今回のデータに基づいて学年全体の平均タイムが *12.5* 秒以下であるかどうかを検 定したい。以下の設問に答えなさい。

   (a) 検定問題として定式化しなさい。

   (b) 検定に用いる分布図を描き、図の中に棄却域と限界値を書き入れたうえで、検定 の結論を述べなさい。

(c) 今回の結論が判断ミスである可能性を確率で示しなさい。この判断ミスは第何種 の過誤に該当するかを言いそえること。

問題 119 ある高校の 1 年生に在学する生徒から無作為に 9 名を選んで 100 メートル走の記録を とったところ

```
12.32,  15.28,  14.19,  13.72,  13.26
14.08,  14.06,  11.82,  12.80,
```

となった。集計計算を行い $\sum_{i=1}^{9} X_i = 121.53,\ \sum_{i=1}^{9} X_i^2 = 1650.165$ という結果まで得た。以下の設問に解答しなさい。

1. 1 年生全体の標準偏差が 1.8 秒だとする。このとき、学年全体の平均タイムが 12.5 秒であると想定するのは無理だろうか。有意水準を 5% として仮説検 定をしなさい。解答するときは帰無仮説 H_0 と対立仮説 H_1 を明記したうえで、途中の計算式も含めて述べること。
2. 1 年生全体の標準偏差が未知として、学年平均 μ を区間推定しなさ い。但し、信頼係数を 95% とする。途中計算も分かりやすく記すこと。

問題 120 5 個の無作為データ

```
1.11, 0.27,  0.81,  0.08,  0.63
```

がとられた分布は標準正規分布 $N(0,1)$ であるか、$\mu > 0$ を母平均とする 別の正規分布 $N(\mu,1)$ であるかを判断したい。以下の設問に解答しなさい。

1. 統計的仮説検定の手順にしたがって検定を行うとすれば帰無仮説 H_0 と 対立仮説 H_1 をどのように定式化すればよいか回答しなさい。
2. 仮説検定の結果がどうなるかを述べなさい。但し、有意水準を 5% と すること。
3. 「第 1 種の過誤」、「第 2 種の過誤」の意味を図を用いて説明しなさい。 但し、対立仮説を $\mu = 0.5$ と置いてよい。
4. 設問 (3) に関連して、対立仮説を $\mu = 0.5$ としたときの検出力を求 めなさい。

問題 121 ブレーキ性能の抜き取り検査をする。まず 10 台の自動車をランダムにとり時速 60Km から急ブレーキをかけて設計どおり 60 メートルで停止できるか どうかを検査する。但し、毎回の停止距離には標準偏差 1 メートルの ばらつきがある。設問に解答しなさい。解答に至る途中のプロセスも分かりやすく書くこと。 ブレーキが甘くなっているかどうかを検査するためには限界値をいくらに定め てどのような検査をすれば良いか。有意水準は 5% とする。 いま工場内に異常があり、平均停止距離は 61 メートルになっている。この検査体 制の検出力は何% ですか。

問題 122 （統計的仮説）検定とは基本的にはデータに基づく二択問題である。 二つある選択肢のうち最初に「正しい」と仮定される仮説を $\boxed{\quad}$ 仮説と言う。データの示す結果から $\boxed{\quad}$ 仮 説が否定されるときは $\boxed{\quad}$ 仮説が採られる。但し、 $\boxed{\quad}$ が真であるときでも誤ってそれを棄却することがある。この種の誤りを $\boxed{\quad}$ の過誤という。他方、 $\boxed{\quad}$ 仮説が正しくてもデータはそのことを示さないことがある。このよう なときは $\boxed{\quad}$ の過誤をおかすことになる。 $\boxed{\quad}$ の過誤を避けるためには $\boxed{\quad}$ を高くする必要がある。

問題 123 ある自動車メーカーではブレーキ性能について、「時速 *60km/h* から急ブレーキをか けた際に停止するまでの距離（＝停止距離）が *50* メートルであるこ と」と定めている。 過去の経験から、実際に測定される停止距離は一様ではなく、正規 分布にしたがっており、標準偏差が *2* メートルであることが確認され ている。ある日、このメーカーで生産されている自動車のブレーキ 性能を検査するため、無作為に *20* 台の自動車を抜き取り、走行テス トを行うことにした。*20* 台を *1* 回ずつ走らせ、時速 *60km/h* に達してか ら、所定の位置で急ブレー キを踏み、停止距離を測定するものとする。以下の設問に解答しなさい。計算プロセスも詳しく述べること。

1. 走行テストを行う自動車のブレーキがすべて正常であると仮定して、停止距離が *49* メートルから *51* メートルまでの範 囲におさまるものは何台程度あると見込まれますか。概算で解答 しなさい。
2. *20* 台の停止距離の平均値を $\bar{X}$ と する。ブレーキ性能に異常がないと仮定し、$\bar{X}$ の標本分布の形を図に描 きなさい。その分布の平均と標準偏差を図の中に書き入れなさい。
3. ブレーキ性能が所定の基準に達していないと判断されれば、生産ラインの一斉点 検を行う予定である。どのような検定問題をとりあげれば、一斉 点検の有無について適切な判断を行えるか。帰無仮説（H_0）、対立 仮説（H_1）を並べて、検定問題を定式化しなさい。新し い記号を使う際は、その記号の意味を記しておくこと。
4. *20* 台の走行テストの結果、$\bar{X}$ は *51* メートルとなった。一斉点検は行うべ きかどうかを回答しなさい。但し、有意水準は *5%* とする。判断を 行った根拠も記述すること。
5. 実際に生産ラインに異常があり、ブレーキ性能が所定の基準に満たない「平均 *53* メートル」になっているものとする。*20* 台の走行テストからこの 異常が認められ、一斉点検に踏み切る確率はどのくらいあります か。但し、有意水準は *5%* とし、停止距離の標準偏差に変化はない ものとします。

問題 124 ある蛍光灯の生産工程から生産される蛍光灯の寿命は過去の記録から平均 *1500* 時 間、標準偏差が *100* 時間の正規分布にしたがっていることがわかっている。ある日、この生産工程から無作為に *50* 本の蛍光灯を抜き出し、 その寿命の標本調査を行うことになった。以下の設問に解答しなさ い。計算プロセスも詳しく述べること。

1. 抜き出した標本の中で寿命が *1700* 時間を超えるものは何本程度あると見込まれる か。概算で解答しなさい。
2. 抜き出した標本の平均寿命を $\bar{X}$ とする。$\bar{X}$ の標本分布を図に描き、 その分布の平均値と標準偏差を図の中に書き込みなさい。
3. この生産工程で生産される蛍光灯の平均寿命が *1500* 時間に達していないことが抜 き出した標本から確認されれば生産を停め一斉点検をしたい。ど のような検定問題をとりあげるのが適当か。帰無仮説（H_0）と対立 仮説（H_1）を並べて、問題を定式化しなさい。但し、新 たな記号を用いる時には、それが何を意味するかを記すこと。
4. 有意水準を *5%* として棄却域を定めるとすると、どのような区間を棄却域とする のが適当ですか。また、標本調査の結果 $\bar{X}$ が棄却域に入っ た時と入らない時のそれぞれについて、一斉点検の可否を述べなさい。
5. この生産工程に実際に障害があり、生産されている蛍光灯の平均寿命が *1470* 時間 に低下しているものとする。標本調査の結果から、この異常に気 がつき、一斉点検に踏み切ることができる確率はどの程度ありま すか。但し、蛍光灯の寿命の標準偏差に変化は生じていないもの とします。

問題 125 ある美容室が割引サービスを行った、この割引サービスによって、一日の平均来客数が増えたかどうかを調べたい。この美容室の普段の平均来客数が 10 人。割引サービスを実施後、25 日間来客数を集計して平均と標準偏差を計算して $\bar{X}=12$、$s=3$ だとする。検定を行って一日平均の来客数が増えたかどうかを判断してください。ヒント：H_0：$\mu \leq 10$；H_1：$\mu > 10$。

問題 126 ある組立工場が大量の部品を入荷した。この種の部品の規格は直径 10 cm である。標本サイズが 100 の無作為標本を抽出して、部品の直径を計り、計算した結果、標本平均 $\bar{X}=11$、標本標準偏差 $S=3$ となった。5% の有意水準で、入荷した部品の直径の平均 μ が 10 cm であるかどうかを検定しなさい。ただし、自由度 99 の t 分布の右側 2.5% 臨界値が 2.276 である。

問題 127 ある食品安全検査機関が食品の発ガン性に関して統計検定を行う。帰無仮説は発ガン性がないとする。有意水準を 5% から 1% へ変更したとする。第 1 種過誤や第 2 種の過誤の意味合いを考えてこの変更は消費者の安全にどういう影響を与えるかについて述べなさい。帰無仮説は発ガン性があるとする場合は？

問題 128 派遣社員の待遇に関して男女差別を調べるために、派遣社員男性と女性 100 人ずつ無作為標本を得たとする。標本に関して月平均所得が男性 23 万円、女性 21 万円だったとする。標本分散が男性 9, 女性 4 とする。両側検定を行いなさい。

問題 129 職業訓練の効果があるかどうかを調べたい。職業訓練を受けたグループと受けていないグループからそれぞれ 50 人ずつ抽出したとする。職業訓練を受けたグループの標本の平均所得 $\bar{X}_1=300$、標本分散 $S_1^2=20$、受けていないグループの標本平均 $\bar{X}_2=290$、標本分散 $S_2^2=18$ とする。ただし、二つのグループの母分散が等しいとする。検定統計量の公式は以下の通りである。5% の有意水準で 2 つのグループの平均所得が同じであるかどうかを検定しなさい。

$$Z=\frac{\left(\bar{X}_1-\bar{X}_2\right)/\sqrt{1/n_1+1/n_2}}{\sqrt{\frac{1}{n_1+n_2-2}\left\{\left(n_1-1\right)S_1^2+\left(n_2-1\right)S_2^2\right\}}}$$

検定するとき、Z の自由度を答えて、以下の二つの有意水準点をどれか正しいものを使ってください。自由度 99 の t 分布の右側 2.5% 有意水準点は 2.2760、自由度 98 の t 分布の右側 2.5% 有意水準点は 2.2763 である。