確率・統計　宿題2　解答例

1.　$\{X_i\} \sim i.i.d.N(\mu, \sigma^2)$ より $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \frac{\sigma^2}{n})$.
$Z = \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$ であるから，$Z^2 = \frac{n(\bar{X}-\mu)^2}{\sigma^2} \sim \chi^2(1)$.
また，$P(Z^2 < a) = P(-\sqrt{a} < Z < \sqrt{a})$（$a > 0$）なので，自由度1の χ^2 分布に関する確率は，標準正規分布に対する確率から計算できる．

2.　$t = \sqrt{n}(\bar{X} - \mu)/S = [\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}]/\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}$.
（分子）：$\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$.（分母）：$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.
また，$\bar{X}$ と S^2 は互いに独立なので，分子，分母は独立である．したがって t 分布の定義より，$t = \sqrt{n}(\bar{X} - \mu)/S$ は，自由度 $n-1$ の t 分布に従う．

3.　$W = \sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} W_i$．ただし $Z_i \sim i.i.d.N(0,1)$，$W_i = Z_i^2 \sim i.i.d.\chi^2(1)$，また，$E(W_i) = 1$，$V(W_i) = 2$．ここで，$\bar{W} = \frac{1}{n}\sum_{i=1}^{n} W_i$，$(i = 1, \cdots, n)$ とすると，$(W-n)/\sqrt{2n} = \sqrt{n}(\bar{W}-1)/\sqrt{2}$．したがって中心極限定理より，$(W-n)/\sqrt{2n} = \sqrt{n}(\bar{W}-1)/\sqrt{2} \xrightarrow{d} N(0,1)$ が成り立つ．

4.　(1) 真である．（$\because$ パラメータ p のベルヌーイ分布からのランダム標本を $\{X_i\}$, $i = 1, \cdots, n$ とすると，$X/n = \frac{1}{n}\sum_{i=1}^{n} X_i$ は平均 p，分散 $p(1-p)$ のベルヌーイ分布からの標本平均である．よって大数の法則より $plim_{n\to\infty} X/n = plim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} X_i = p$.）or（$\because$ $X \sim B(n,p)$ であるから, $E(X/n) = p$, $V(X/n) = p(1-p)$．このとき，$\forall\varepsilon > 0$，$P(|X/n - p| > \varepsilon) = P((X/n-p)^2 > \varepsilon^2) \leq p(1-p)/n\varepsilon^2 \to 0$ as $n \to \infty$(不等号はチェビシェフの不等式より)．したがって，$n \to \infty$ のとき，X/n は p に確率収束する．）
(2) 偽である．（$\because$ X は標本平均ではないので，大数の法則は適用できず，確率収束をいうことはできない．）or（$\because$ 中心極限定理より $X \xrightarrow{d} N(np, np(1-p))$ であるから，任意の正数 ε に対して $P(|X - np| > \varepsilon) = P(\frac{|X-np|}{\sqrt{np(1-p)}} > \frac{\varepsilon}{\sqrt{np(1-p)}}) = P(|Z| > \frac{\varepsilon}{\sqrt{np(1-p)}}) \overset{n\to\infty}{\to} P(|Z| > 0) = 1$．したがって，$X$ と np の差は $n \to \infty$ のとき0に確率収束するとはいえない．）

5.　$E(X_i) = 1/\lambda$，$V(X_i) = 1/\lambda^2$ であるから，$E(\bar{X}) = 1/\lambda$，$V(\bar{X}) = 1/n\lambda^2$．したがって，$\bar{X} \approx N(1/\lambda, 1/n\lambda^2)$.
このとき，$P(\bar{X} > a) = P(\frac{\bar{X}-1/\lambda}{\sqrt{1/n\lambda^2}} > \frac{a-1/\lambda}{\sqrt{1/n\lambda^2}}) = P(Z > \sqrt{n}(\lambda a - 1))$.

6.　一様分布 $U(-50,50)$ に従う確率変数 X_i $(i.i.d., i=1,\cdots,100)$ の和の絶対値が 200 を超える確率を正規近似で求めればよい．

$E(X_i)=0$，$V(X_i)=100^2/12=2500/3$ であるから，$Y=\sum_{i=1}^{100}X_i$ について，$E(Y)=0$，$V(Y)=25\times10^4/3$．したがって，$Y\approx N(0,25\times10^4/3)$．

$P(|Y|>200)=P(|Y|/\sqrt{25\times10^4/3}>200/\sqrt{25\times10^4/3})=P(|Z|>0.69)=2P(Z>0.69)=2\times0.2451=0.4902$.

7.　第 i 日の来客数 X_i は他の日と独立にパラメータ 10 のポワソン分布に従うと考えると，$E(X_i)=V(X_i)=10$．したがって，30 日間の来客数 $Y=\sum_{i=1}^{30}X_i$ について，$E(Y)=30\times10=300$，$V(Y)=30\times10=300$，$Y\approx N(300,300)$.

これより，$P(Y\geq400)=P((Y-300)/\sqrt{300}\geq(400-300)/\sqrt{300})=P(Z\geq10/\sqrt{3})=P(Z\geq5.7735)=0.00000$ であるから，この売り場に 30 日間で 400 人以上の来客がある確率は，ほぼ 0 である．

8.　$\hat{\alpha}$ について：期待値 $E(\hat{\alpha})=E(2\overline{X})=2\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{2}{n}\frac{n\alpha}{2}=\alpha$．したがって $\hat{\alpha}$ は不偏推定量である．また，分散 $V(\hat{\alpha})=V(2\overline{X})=4V(\overline{X})=4\alpha^2/12n=\alpha^2/3n$.

$\tilde{\alpha}$ について：まず, $X_1,\cdots,X_n$ を小さいほうから順に並べたものを $X_{(1)},\cdots,X_{(n)}$ と表すと, $\tilde{\alpha}=\frac{n+1}{n}X_{(1)}$．$X_{(1)}$ の密度関数は $f_{(1)}(x)={}_nC_1f(x)[1-F(x)]^{n-1}=n(-1/\alpha)[1-(1-x/\alpha)]^{n-1}=-nx^{n-1}/\alpha^n$ （$f(x)$ は X_i の密度関数で $f(x)=-1/\alpha$，$F(x)$ は分布関数で $F(x)=\int_{\alpha}^{x}(-1/\alpha)du=1-x/\alpha$.）であるから，$E(\tilde{\alpha})=E(\frac{n+1}{n}X_{(1)})=\int_{\alpha}^{0}\frac{n+1}{n}x(-x^{n-1}/\alpha^n)dx=-\frac{n+1}{\alpha^n}\int_{\alpha}^{0}x^ndx=(-1/\alpha^n)\times(-\alpha^{n+1})=\alpha$．したがって，$\tilde{\alpha}$ は不偏推定量である．また，分散 $V(\tilde{\alpha})=V(\frac{n+1}{n}X_{(1)})=\frac{(n+1)^2}{n^2}V(X_{(1)})$．$V(X_{(1)})=\int_{\alpha}^{0}x^2(-nx^{n-1}/\alpha^n)dx-(\frac{n}{n+1}\alpha)^2=-\frac{n}{\alpha^n}\int_{\alpha}^{0}x^{n+1}dx-(\frac{n}{n+1}\alpha)^2=\frac{n}{n+2}\alpha^2-(\frac{n}{n+1}\alpha)=n\alpha^2/(n+1)^2(n+2)$．$\therefore V(\tilde{\alpha})=\frac{(n+1)^2}{n^2}V(X_{(1)})=\alpha^2/n(n+2)\leq\alpha^2/3n=V(\hat{\alpha})$.

推定量の優劣について：$V(\hat{\alpha})\geq V(\tilde{\alpha})$ であるから，$\tilde{\alpha}$ の方が優れた不偏推定量である．

一様分布 $U(0,\beta)$ からのランダム標本 $Y_1,\cdots,Y_n$ について：$\beta=-\alpha$, $Y_i=-X_i$ と考えればよい．すると，$-X_{(1)}=Y_{(n)}$，$\cdots$，$-X_{(n)}=Y_{(1)}$（ただし，$Y_{(j)}$ は $Y_1,\cdots,Y_n$ の中で j 番目に大きいものを表す.），$f(X_{(1)})=f(Y_{(n)})$，$\cdots$，$f(X_{(n)})=f(Y_{(1)})$．したがって $\beta=-\alpha$ の 2 つの不偏推定量は, 先の議論を利用して $\hat{\beta}=-\hat{\alpha}=-2\overline{X}=2\overline{Y}$ および $\tilde{\beta}=-\tilde{\alpha}=-\frac{n+1}{n}X_{(1)}=\frac{n+1}{n}Y_{(n)}=\frac{n+1}{n}max(Y_1,\cdots,Y_n)$．分散はそれぞれ $V(\hat{\beta})=\beta^2/3n$, $V(\tilde{\beta})=\beta^2/n(n+2)$ なので，$V(\hat{\beta})\geq V(\tilde{\beta})$．したがって $\tilde{\beta}$ の方が優れた推定量である．