Data provided: Formulae and Critical Values Sheet

MAS6012

The University Of Sheffield.

SCHOOL OF MATHEMATICS AND STATISTICS **Spring Semester 2016–2017**

Sampling, Design, Medical Statistics **3 hours**

Candidates may bring to the examination a calculator that conforms to University regulations. All answers will be marked but credit will be given only for the best ***FIVE*** *answers. All questions are worth 20 marks. Total marks 100.*

Please leave this exam paper on your desk
Do not remove it from the hall

Registration number from U-Card (9 digits)
to be completed by student

Blank

1 The data below come from an RCT of a possible new drug for treatment of hypertension in diabetics (adapted from Altman (1995); original data from Hommel *et al.* (1986)). Systolic blood pressures of 16 patients before and after one week's treatment with Captopril or Placebo are given in mmHg.

	Captopril				Placebo		
	Baseline	After 1 week	Change		Baseline	After 1 week	Change
1	147	137	-10	1	133	139	6
2	129	120	-9	2	129	134	5
3	158	141	-17	3	152	136	-16
4	164	137	-27	4	161	151	-10
5	134	140	6	5	154	147	-7
6	155	144	-11	6	141	137	-4
7	151	134	-17	7	156	149	-7
8	141	123	-18				
9	153	142	-11				
Mean	148.0	135.3	-12.7	Mean	146.6	141.9	-4.7
SD	11.43	8.43	8.99	SD	12.29	6.94	7.91

(i) Explain what is meant by an RCT and what advantages they have.
(4 marks)

(ii) The authors of the original study performed paired t tests on the data in each group. They found a significant change in pressure in the Captopril group (t=4.24, df=8, p=0.003), but not in the Placebo group (t=1.57, df=6, p=0.17). They then concluded that therefore 'Captopril represents a valuable new drug for treating hypertension in diabetics'.

(a) What is wrong with their analysis? ***(3 marks)***

(b) Suggest a better analysis, explaining your reasoning. Perform the test you suggest and explain what conclusions can be drawn about the effects of Captopril in such patients. ***(8 marks)***

(iii) In the trial the group sizes are unbalanced (Captopril:Placebo is 9:7 rather than 8:8).

(a) What effect is this likely to have on the effectiveness of the trial?
(2 marks)

(b) Specify a mechanism by which one could ensure a balanced allocation of the 16 patients. Give details of how it could be applied.
(3 marks)

2 The data in parts (i) and (ii) of this question come from a study of adverse reactions in recipients of bone marrow transplants for treatment of leukaemia (adapted from Altman (1995); original data from Bagot *et al.* (1988)). The study concerned 37 transplants in which 17 recipients went on to develop an adverse reaction while 20 did not. Possible explanatory factors which were also recorded were: type of leukaemia (coded AML, ALL or CML), recipient and donor ages (in years), whether the donor had ever been pregnant (coded 1= Yes, 0=No) and a continuous chemical index (Index) reflecting condition at the time of transplant.

(i) Use the data below to assess whether type of leukaemia is associated with occurrence of an adverse reaction.

	Type of leukaemia		
	AML	ALL	CML
Reaction	5	4	8
No reaction	6	12	2

(4 marks)

(ii) A logistic regression analysis of the full data set, using the stepwise method and a 5% level of significance to assess whether to retain a variable, gave the output below. Here the response has been coded 1=adverse reaction occurred, 0=no adverse reaction occurred; leukaemia type was recoded by use of 2-level dummy variables for presence of each of ALL and CML; and only main effects were considered.

Variable	Coefficient	SE	z	p-value
CML	2.251	1.106	2.035	0.04
Pregnancy	2.496	1.101	2.266	0.02
Index	1.488	0.720	2.067	0.04

(a) Explain why a logistic regression analysis might be preferable to analysis of the type in part (a). ***(2 marks)***

(b) Use the output above to calculate the probability of an adverse reaction for a patient age 20, with ALL and an Index of 0.9, receiving a donation from a donor age 25 who had been pregnant. ***(4 marks)***

2 (continued)

(iii) An engineering firm is testing a new method of making drill bits (the cutting tool used on a drill). In a trial, 9 drill bits were tested to destruction. Follow-up was censored prior to failure in 3 of the bits. The table below shows the times until failure in weeks. An asterisk denotes a censored observation.

	New Time (weeks)
	1
	4
	11*
	15*
	16
	32
	34
	52*
	85
Total	250

(a) Assuming that the censoring is for reasons unrelated to drill bit breakage, estimate the value of the survivor function at 20 weeks using a Kaplan-Meier estimate. ***(4 marks)***

Note: You are not required to work out the Kaplan-Meier estimates beyond this time.

(b) It is suggested that the survival times are exponentially distributed. Under this assumption, estimate the value of the survivor function at 20 weeks. ***(3 marks)***

Hint: For an exponential distribution with rate λ, then $S(t) = e^{-\lambda t}$.

(c) How would you decide if the exponential distribution assumption in part (b) was appropriate? ***(1 mark)***

(d) If censoring in this study was due to the operator noticing a hairline crack in the drill bit, briefly discuss the appropriateness of using the above failure times to assess the success of the new manufacturing technique. ***(2 marks)***

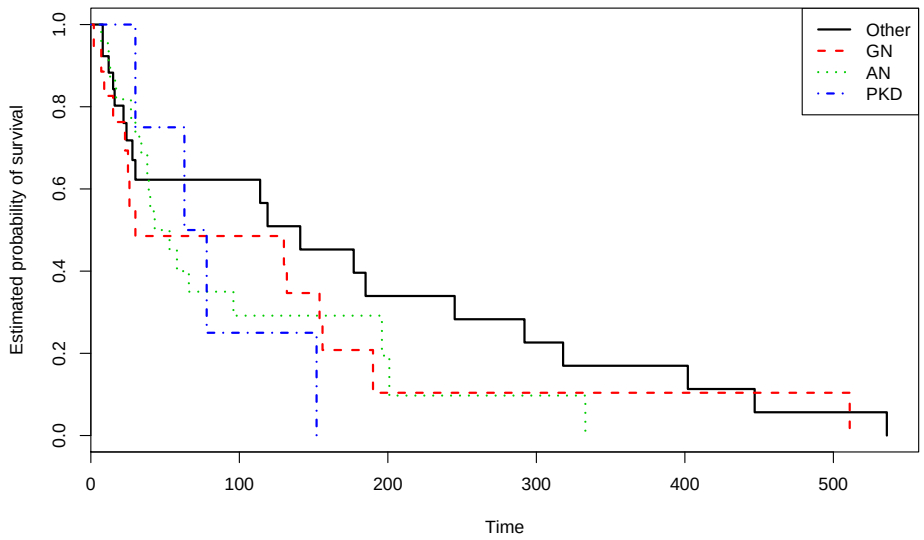
3 McGilchrist and Aisbett (1991) reported a trial on kidney disease patients using portable dialysis equipment. When an infection is found in a dialysis patient, the catheter they currently have needs to be removed and the infection cleared up. Once this is done a new catheter is inserted. McGilchrist and Aisbett investigated the recurrence time to infection, from the time of insertion of this new catheter. Catheters were removed for reasons other than infection, in which case the observation was right censored. Primary interest was in whether the disease type affected recurrence time.

The data are stored in `dialysis` and coding for the different variables is shown below:

Coding:
time: time until recurrence (weeks)
status: censoring indicator (1 = recurrence; 0 = censoring)
age: age of patient (yrs)
sex: 0 = Female, 1 = Male
disease type, four possible types: GN, AN, PKD, and Other

(i) Briefly describe what procedure the following R code implements and what the plot that it produces tells you.

```
> attach(dialysis)
> dialysis.surv <- Surv(time, status)
>
> dialysisfit.raw <- survfit(dialysis.surv ~ disease)
> plot(dialysisfit.raw, col=c(1:4), lty=c(1:4),lwd=2,xlab="Time",
+   ylab="Estimated probability of survival")
> legend("topright",levels(disease),col=c(1:4),lty=c(1:4),lwd=2)
```



(4 marks)

3 (continued)

(ii) A more in depth analysis was performed below:

```
> dialysis.fit <- survreg(dialysis.surv ~ sex + age + disease,
+   dist = "exponential")
> summary(dialysis.fit)

Call:
survreg(formula = dialysis.surv ~ sex + age + disease, dist = "exponential")
               Value Std. Error      z        p
(Intercept)  5.20601     0.4126 12.617 1.71e-36
sex         -1.74338     0.3126 -5.578 2.44e-08
age          0.00235     0.0112  0.210 8.34e-01
diseaseGN   -0.08935     0.4040 -0.221 8.25e-01
diseaseAN   -0.59722     0.3934 -1.518 1.29e-01
diseasePKD  -0.21870     0.6424 -0.340 7.34e-01

Scale fixed at 1

Exponential distribution
Loglik(model)= -312.4   Loglik(intercept only)= -326.5
Chisq= 28.07 on 5 degrees of freedom, p= 3.5e-05
Number of Newton-Raphson Iterations: 5
n= 74

>
> anova(dialysis.fit)
        Df   Deviance Resid. Df    -2*LL     Pr(>Chi)
NULL    NA         NA        73 652.9310           NA
sex      1 24.9099653        72 628.0211 6.007110e-07
age      1  0.2971188        71 627.7239 5.856943e-01
disease  3  2.8642543        68 624.8597 4.130315e-01
```

Describe this analysis making sure you explain:

(a) The type of model fitted; the statistical representation of the model; and the coding of the variables; ***(5 marks)***

(b) What the effects of the different variables are on recurrence; and if disease type is seen to make a difference. ***(5 marks)***

(iii) Using the model output in part (ii), predict the expected recurrence time for a 40yr old male with disease type 'Other'. ***(4 marks)***

(iv) Write down one benefit to a clinician that this model provides that a proportional hazards model would not. Similarly write down one benefit of a proportional hazards model. ***(2 marks)***

4 (i) A trial is designed to compare the effectiveness of three drugs numbered $1, 2$ and 3. There are nine participants in total in the trial. Three participants are given Drug 1, three are given Drug 2 and three and given Drug 3. The following model is proposed

$$EY_{ij} = \mu + \alpha_i,$$

where Y_{ij} is the response of the jth participant given drug i for $1 \leq i, j \leq 3$.

(a) Write down the design matrix $\boldsymbol{X}$ for this design with parameters $\mu, \alpha_1, \alpha_2, \alpha_3$ and explain what is wrong with this parameterisation. ***(3 marks)***

(b) By applying the constraint $\alpha_1 + \alpha_2 + \alpha_3 = 0$ write down the design matrix with parameters μ, α_1, α_2. ***(2 marks)***

(c) Suppose instead that there are m participants in total of which t are given the Drug 1, t are given Drug 2 and $m - 2t$ are given Drug 3. For the model $EY_{ij} = \mu + \alpha_i$ for $i, j = 1, 2, 3$ with constraint $\alpha_1 + \alpha_2 + \alpha_3 = 0$ find $\boldsymbol{X}^T\boldsymbol{X}$ in terms of m and t, where $\boldsymbol{X}$ is the design matrix. ***(4 marks)***

(d) Given that $|\boldsymbol{X}^T\boldsymbol{X}| = 9t^2(m - 2t)$ find the integer value of t that gives a $D-$optimal design for $m = 1000$. ***(5 marks)***

(ii) Consider a fractional factorial design with 4 factors (x_1, x_2, x_3, x_4) each of which occurs at two levels, denoted $+1$ and -1.

(a) Suppose that four design points are available. Provide two design generators that allow the intercept and the main effects for x_1, x_2 and x_4 to be included in the linear model without confounding. Show the alias structure for these two generators. ***(3 marks)***

(b) Construct the fractional factorial design using your design in part (ii)(a). ***(3 marks)***

5 An investigator is studying the dependence of a variable Y on a continuous explanatory variable x which has been scaled to lie between -1 and 1. Each observation is independently subject to a measurement error with mean 0 and variance σ^2. The following model is proposed for the ith observation

$$EY_i = \beta_0 + \beta_1 x_i^3.$$

The investigator proposes to use the four design points $\{x_1, x_2, x_3, x_4\} = \{-1, 0, 1, 1\}$. Denote the response for these design points Y_1, Y_2, Y_3, Y_4 respectively.

(i) If $\boldsymbol{X}$ is the design matrix show that

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1} = \frac{1}{11}\begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix}$$

and hence give the variances of the least squares estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ in terms of σ^2 for this design. ***(3 marks)***

(ii) The investigator thinks the 95% confidence region for $(\hat{\beta}_0, \hat{\beta}_1)^T$ will be an ellipse (rather than a circle) with major and minor axes parallel to the co-ordinate axes. Justify whether they are correct. ***(2 marks)***

(iii) Derive the cartesian co-ordinates of the centre of the 95% confidence region for $(\hat{\beta}_0, \hat{\beta}_1)^T$ in terms of Y_1, Y_2, Y_3, Y_4. ***(4 marks)***

(iv) Show that for the model $EY_i = \beta_0 + \beta_1 x_i^3$, the design $\{x_1, x_2, x_3, x_4\} = \{-1, 0, 1, 1\}$ is neither D-optimal nor G-optimal, by using the General Equivalence Theorem. ***(6 marks)***

(v) Suppose a design point $\boldsymbol{x_0}$ is to be removed from a design ξ with information matrix $\boldsymbol{G}$. Let $\boldsymbol{G}^*$ be the information matrix of the design ξ with $\boldsymbol{x_0}$ removed and $\boldsymbol{f}(\boldsymbol{x_0})^T$ be the row of the design matrix for ξ corresponding to point $\boldsymbol{x_0}$. Show that $|\boldsymbol{G}^*| = |\boldsymbol{G}|\left(1 - \boldsymbol{f}(\boldsymbol{x_0})^T\boldsymbol{G}^{-1}\boldsymbol{f}(\boldsymbol{x_0})\right)$. You may find the following result useful.

If $\boldsymbol{A}$ is a non-singular $p \times p$ matrix, $\boldsymbol{B}$ is a $p \times n$ matrix, $\boldsymbol{C}$ is a $n \times p$ matrix and $\boldsymbol{I}$ is the $n \times n$ identity matrix, then $|\boldsymbol{A} - \boldsymbol{BC}| = |\boldsymbol{A}||\boldsymbol{I} - \boldsymbol{CA}^{-1}\boldsymbol{B}|$.
(3 marks)

(vi) The investigator is to remove a point from the current design $\{x_1, x_2, x_3, x_4\} = \{-1, 0, 1, 1\}$ for the model $EY_i = \beta_0 + \beta_1 x_i^3$. Justify the D-optimal choice of point to remove from this design. ***(2 marks)***

6 A survey conducted to estimate the mean annual spend on organic food by postgraduate students at Sheffield University produced the following data

Stratum	Population size	Sample size	std. dev. (£)	mean (£)
1	1000	30	32	450
2	2000	50	20	300

(i) Estimate the mean annual spend on organic food by postgraduate students at Sheffield University using the best linear unbiased estimator $\bar{x}_{st}$.
(2 marks)

(ii) Estimate a 95% confidence interval for the population mean using $\bar{x}_{st}$. State any assumptions you make. ***(5 marks)***

(iii) Another survey is to be conducted with the same strata. If the sampling costs are such that the total sample size is now 250, specify the sample sizes in each stratum using each of the following methods

(a) Neyman allocation; ***(3 marks)***

(b) minimising the variance of $\bar{x}_{st}$, subject to a fixed total cost but unequal strata sampling costs: a student from stratum 2 is twice as expensive to sample as a student from stratum 1. ***(4 marks)***

(iv) Suppose the surveyor chose the stratum sample sizes by minimising the variance of $\bar{x}_{st}$, subject to a fixed total cost but decided that a student from stratum 2 was k times as expensive to sample as a student from stratum 1. If the surveyor allocated equal sample sizes to strata 1 and 2, what value of k did they assume? ***(2 marks)***

(v) Suppose instead that simple random sampling is to be used. A pilot study using simple random sampling surveyed 10 students and gave the following data (where x_i is the ith Sheffield postgraduate student annual spend on organic food in £)

$$\sum_{i=1}^{10} x_i = 2800 \quad \sum_{i=1}^{10} x_i^2 = 1,110,250$$

Using the data from the pilot study what sample size is needed for the width of the 95% confidence interval for the annual spend on organic food to be no more than £5. ***(4 marks)***

End of Question Paper

MAS6012 FORMULA SHEET & CRTICAL VALUES

1 Clinical Trials Formulae

Two Sample t-Test — Separate variances form $r = min(n_1, n_2)$

$$t_r \;=\; |\frac{\bar{X_1} - \bar{X_2}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}|$$

Two Sample t-Test — Pooled variance form $r = n_1 + n_2 - 2$

$$t_r \;=\; |\frac{\bar{X_1} - \bar{X_2}}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}(\frac{1}{n_1} + \frac{1}{n_2})}}|$$

Sample Size Calculations — Two sample test for proportions NB number in each group

$$n \simeq \frac{\theta_2(1-\theta_2) + \theta_1(1-\theta_1)}{(\theta_2 - \theta_1)^2}[\Phi^{-1}(\beta) + \Phi^{-1}(\alpha/2)]^2$$

Sample Size Calculations — Two sample test for means NB number in each group

$$n \simeq \frac{2\sigma^2}{(\mu_2 - \mu_1)^2}[\Phi^{-1}(\beta) + \Phi^{-1}(\alpha/2)]^2$$

Standard Error for Natural Logarithm of Relative Risk

$$s.e[(log_e(RR)] = \sqrt{\frac{1}{a} - \frac{1}{a+b} + \frac{1}{c} - \frac{1}{c+d}}$$

Standard Error for Natural Logarithm of Odds Ratio

$$s.e[(log_e(OR)] = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

2 Survival Analysis Formulae

Exponential Distributions — MLE of rate λ with censoring The mle

$$\hat{\lambda} = \frac{\sum_{i=1}^n \delta_i}{\sum_{i=1}^n t_i} = \frac{\Delta}{\mathcal{T}} \qquad var(\hat{\lambda}) \approx \frac{\hat{\lambda}^2}{\sum_{i=1}^n \delta_i}.$$

For any (differentiable, monotonic) function $g(\cdot)$,

$$var(g(\hat{\lambda})) \approx \left[\{g'(\lambda)\}^2 var(\lambda)\right]_{\lambda=\hat{\lambda}}$$

so e.g.

$$var(\frac{1}{\hat{\lambda}}) = var(\hat{\mu}) \approx \frac{\hat{\mu}^2}{\sum_{i=1}^n \delta_i}$$

Exponential Distributions — MLE test

$$W = \frac{\hat{\lambda}_1 - \hat{\lambda}_2}{\sqrt{\frac{\hat{\lambda}_1^2}{\Delta_1} + \frac{\hat{\lambda}_2^2}{\Delta_2}}} \approx N(0,1)$$

Exponential Distributions — LRT test

$$2\left\{\Delta_1 \log\frac{\Delta_1}{\mathcal{T}_1} + \Delta_2 \log\frac{\Delta_2}{\mathcal{T}_2} - (\Delta_1 + \Delta_2)\log\frac{\Delta_1 + \Delta_2}{\mathcal{T}_1 + \mathcal{T}_2}\right\} \approx \chi_1^2$$

Log-rank Statistic

$$LR = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} \sim \chi_1^2$$

3 Design Formulae

Linear Model formulae

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \quad \text{and} \quad \hat{\boldsymbol{\beta}} \sim N\{\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}\}$$

Prediction Variance

$$\text{var } \hat{y}(\boldsymbol{x}_0) = \sigma^2 \boldsymbol{f}(\boldsymbol{x}_0)^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{f}(\boldsymbol{x}_0)$$

Standardized Prediction Variance

$$d(\boldsymbol{x}) = n\boldsymbol{f}(\boldsymbol{x})^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{M}^{-1}\boldsymbol{f}(\boldsymbol{x})$$

Confidence Regions, σ^2 unknown

$$p^{-1}\hat{\sigma}^{-2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T\boldsymbol{X}^T\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \text{ has an } F_{p,n-p} \text{ distribution, provided } n > p$$

Balanced Incomplete Block Design Notation

$$\begin{aligned} t &= \text{number of treatments} \\ k &= \text{number of units in a block} \\ b &= \text{number of blocks} \\ r &= \text{number of applications of each treatment} \\ \lambda &= \text{number of times each pair of treatments appears together in a block} \end{aligned}$$

Balanced Incomplete Block Design Relationships

$$\begin{aligned} t &> k \\ bk &= rt \\ r(k-1) &= \lambda(t-1) \end{aligned}$$

Balanced Incomplete Block Design - Unreduced Design

$$b = \binom{t}{k} \quad r = \binom{t-1}{k-1} \quad \lambda = \binom{t-2}{k-2}$$

Efficiency of Balanced Incomplete Block Design compared to a Randomized Block design

$$\frac{1 - t^{-1}}{1 - k^{-1}}$$

Adding an extra point $\boldsymbol{x}$

$$|\boldsymbol{G}^*| = |\boldsymbol{G}|\left(1 + \boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{G}^{-1}\boldsymbol{f}(\boldsymbol{x})\right)$$

Deleting an existing point $\boldsymbol{x}$

$$|\boldsymbol{G}^*| = |\boldsymbol{G}|\left(1 - \boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{G}^{-1}\boldsymbol{f}(\boldsymbol{x})\right)$$

Adding a new point $\boldsymbol{y}$ and deleting an existing point $\boldsymbol{x}$

$$|\boldsymbol{G}_2| = |\boldsymbol{G}|\left\{\left(1 - \boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{G}^{-1}\boldsymbol{f}(\boldsymbol{x})\right)\left(1 + \boldsymbol{f}(\boldsymbol{y})^T\boldsymbol{G}^{-1}\boldsymbol{f}(\boldsymbol{y})\right) + \left(\boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{G}^{-1}\boldsymbol{f}(\boldsymbol{y})\right)^2\right\}$$

4 Sample Surveys and Computer Experiments Formulae

Population variance

$$S^2 = \frac{1}{N-1}\sum_1^N (X_i - \overline{X})^2 = \frac{1}{N-1}\left(\sum_{i=1}^N X_i^2 - N\bar{X}^2\right)$$

and for a binary characteristic ($X_i = 1$ or 0 for each i),

$$S^2 = \frac{NP(1-P)}{N-1}$$

Variance of sample mean for simple random sampling

$$\text{var } \overline{x} = \left(1 - \frac{n}{N}\right)\frac{S^2}{n}.$$

Sample size to achieve given confidence interval width for simple random sampling

$$n \geq \frac{N}{1 + N(d/(2Sz_{\alpha/2}))^2}$$

Stratified estimate of population mean and its variance

$$\overline{x}_{st} = \frac{1}{N}\sum_1^l N_i\overline{x}_i \quad \text{and} \quad \text{var } \overline{x}_{st} = \sum_1^l \left(\frac{N_i}{N}\right)^2 \frac{1-f_i}{n_i}S_i^2.$$

Allocation methods

$$\text{Optimal allocation: } n_i \propto \frac{N_iS_i}{\sqrt{c_i}} \qquad \text{Neyman allocation: } n_i = \frac{nN_iS_i}{\sum_1^l N_iS_i}$$

Cluster estimate of population mean and its variance

$$\overline{x}_{cl} = \frac{1}{lK}\sum_1^l\sum_1^K x_{ij} \quad \text{and} \quad \text{var } (\overline{x}_{cl}) = \frac{1-f}{l}\frac{1}{L-1}\sum_1^L(\overline{X}_i - \overline{X})^2$$

Regression estimator of population mean and its variance

$$\overline{x}_{lr} = \overline{x} - \hat{\beta}(\overline{y} - \overline{Y}) \quad \text{and} \quad \text{var } \overline{x}_{lr} \simeq \frac{1-f}{n}S_X^2(1-\rho^2)$$

Approximate variance of the Peterson estimator, Chapman estimator and approximate variance

n: size of 1st sample, m: size of 2nd sample.

$$\begin{aligned}
\widehat{Var}(\hat{N}_p) &= \frac{mn^2(m-r)}{r^3}, \\
\hat{N}_c &= \frac{(n+1)(m+1)}{r+1} - 1, \\
\widehat{Var}(\hat{N}_c) &= \frac{(n+1)(m+1)(n-r)(m-r)}{(r+1)^2(r+2)}.
\end{aligned}$$

Variance identity

$$Var(Y) = Var_X\{E(Y|X)\} + E_X\{Var(Y|X)\}.$$

5 Tables of Percentage Points (also known as Quantiles or Critical Values) for Three Standard Distributions

The tables contain values of quantiles q such that $P[X \leq q] = p$ for various probabilities p when X has the specified distribution (which may depend on particular degrees of freedom ν). In these tables, p has been expressed as a percentage rather than a decimal. The relevant R commands for generating the q are also shown. For the $N(0,1)$ distribution, the tabulated function is also known as the Φ^{-1} function.

STANDARD NORMAL DISTRIBUTION PERCENTAGE POINTS

`qnorm(`p`)` where p is percentage, e.g. for 95%, $p = 0.95$

	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
`qnorm`	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090

CHI-SQUARED PERCENTAGE POINTS

`qchisq`(p, ν) where p is percentage, e.g. for 95%, $p = 0.95$

ν	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
1	0.708	0.936	1.323	1.642	2.354	2.706	3.841	5.024	6.635	7.879	10.828
2	1.833	2.197	2.773	3.219	4.159	4.605	5.991	7.378	9.210	10.597	13.816
3	2.946	3.405	4.108	4.642	5.739	6.251	7.815	9.348	11.345	12.838	16.266
4	4.045	4.579	5.385	5.989	7.214	7.779	9.488	11.143	13.277	14.860	18.467
5	5.132	5.730	6.626	7.289	8.625	9.236	11.070	12.833	15.086	16.750	20.515
6	6.211	6.867	7.841	8.558	9.992	10.645	12.592	14.449	16.812	18.548	22.458
7	7.283	7.992	9.037	9.803	11.326	12.017	14.067	16.013	18.475	20.278	24.322
8	8.351	9.107	10.219	11.030	12.636	13.362	15.507	17.535	20.090	21.955	26.125
9	9.414	10.215	11.389	12.242	13.926	14.684	16.919	19.023	21.666	23.589	27.877
10	10.473	11.317	12.549	13.442	15.198	15.987	18.307	20.483	23.209	25.188	29.588

STUDENT'S t PERCENTAGE POINTS

`qt`(p, ν) where p is percentage, e.g. for 95%, $p = 0.95$

ν	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
2	0.289	0.500	0.816	1.061	1.604	1.886	2.920	4.303	6.965	9.925	22.327
3	0.277	0.476	0.765	0.978	1.423	1.638	2.353	3.182	4.541	5.841	10.215
4	0.271	0.464	0.741	0.941	1.344	1.533	2.132	2.776	3.747	4.604	7.173
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
6	0.265	0.453	0.718	0.906	1.273	1.440	1.943	2.447	3.143	3.707	5.208
7	0.263	0.449	0.711	0.896	1.254	1.415	1.895	2.365	2.998	3.499	4.785
8	0.262	0.447	0.706	0.889	1.240	1.397	1.860	2.306	2.896	3.355	4.501
9	0.261	0.445	0.703	0.883	1.230	1.383	1.833	2.262	2.821	3.250	4.297
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
11	0.260	0.443	0.697	0.876	1.214	1.363	1.796	2.201	2.718	3.106	4.025
12	0.259	0.442	0.695	0.873	1.209	1.356	1.782	2.179	2.681	3.055	3.930
13	0.259	0.441	0.694	0.870	1.204	1.350	1.771	2.160	2.650	3.012	3.852
14	0.258	0.440	0.692	0.868	1.200	1.345	1.761	2.145	2.624	2.977	3.787
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
16	0.258	0.439	0.690	0.865	1.194	1.337	1.746	2.120	2.583	2.921	3.686
17	0.257	0.438	0.689	0.863	1.191	1.333	1.740	2.110	2.567	2.898	3.646
18	0.257	0.438	0.688	0.862	1.189	1.330	1.734	2.101	2.552	2.878	3.610
19	0.257	0.438	0.688	0.861	1.187	1.328	1.729	2.093	2.539	2.861	3.579
20	0.257	0.437	0.687	0.860	1.185	1.325	1.725	2.086	2.528	2.845	3.552
21	0.257	0.437	0.686	0.859	1.183	1.323	1.721	2.080	2.518	2.831	3.527
22	0.256	0.437	0.686	0.858	1.182	1.321	1.717	2.074	2.508	2.819	3.505
23	0.256	0.436	0.685	0.858	1.180	1.319	1.714	2.069	2.500	2.807	3.485
24	0.256	0.436	0.685	0.857	1.179	1.318	1.711	2.064	2.492	2.797	3.467
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
26	0.256	0.436	0.684	0.856	1.177	1.315	1.706	2.056	2.479	2.779	3.435
27	0.256	0.435	0.684	0.855	1.176	1.314	1.703	2.052	2.473	2.771	3.421
28	0.256	0.435	0.683	0.855	1.175	1.313	1.701	2.048	2.467	2.763	3.408
29	0.256	0.435	0.683	0.854	1.174	1.311	1.699	2.045	2.462	2.756	3.396
30	0.256	0.435	0.683	0.854	1.173	1.310	1.697	2.042	2.457	2.750	3.385
35	0.255	0.434	0.682	0.852	1.170	1.306	1.690	2.030	2.438	2.724	3.340
40	0.255	0.434	0.681	0.851	1.167	1.303	1.684	2.021	2.423	2.704	3.307
45	0.255	0.434	0.680	0.850	1.165	1.301	1.679	2.014	2.412	2.690	3.281
50	0.255	0.433	0.679	0.849	1.164	1.299	1.676	2.009	2.403	2.678	3.261
55	0.255	0.433	0.679	0.848	1.163	1.297	1.673	2.004	2.396	2.668	3.245
60	0.254	0.433	0.679	0.848	1.162	1.296	1.671	2.000	2.390	2.660	3.232
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090