ST3131 Regression Analysis

AY2021/22 Semester 2

1. Simple Linear Regression

1.1. Introduction

- Pearson's Correlation: Let X and Y be two random variables, the theoretical Pearson's Correlation is defined as $\rho_{xy} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\sum_{i=1}^n (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^n (x_i-\bar{x})^2 \sum_{i=1}^n (y_i-\bar{y})^2}}$.
    - $\rho_{xy} > 0 \rightarrow$ positive correlation
    - $\rho_{xy} < 0 \rightarrow$ negative correlation
    - $\rho_{xy} = 0 \rightarrow$ no linear relationship
- Simple Regression Model: $Y = \beta_0 + \beta_1 X + \epsilon$, where β_0, β_1 are regression coefficients and ϵ is a random error.
    - Regression Function: $EY = \beta_0 + \beta_1 X$.
    - General assumptions:
        1) x_i and ϵ_i are independent.
        2) ϵ_i's have mean zero.
        3) ϵ_i's are pairwise uncorrelated ($Cov(\epsilon_i, \epsilon_j) = 0$).
        4) (Homogeneity) ϵ_i's have common variance σ^2.
        5) (Normality) ϵ_i's have a normal distribution.

1.2. The Estimation of LRM

- Least Square Estimation of β_0 and β_1:
  (β_0, β_1) minimizes $E(Y - b_0 - b_1 X)^2$
  $$= \frac{1}{n}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$
  Let $Q = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$. By minimizing Q, we can estimate the parameters:
  $$\widehat{\beta_1} = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{X})^2}$$
  $$\widehat{\beta_0} = \bar{Y} - b_1 \bar{X}$$
  where $\bar{X} = \frac{1}{n}\sum_{i=1}^n x_i$ and $\bar{Y} = \frac{1}{n}\sum_{i=1}^n y_i$.
- Least Square Estimation of σ^2:
  Since $\epsilon \sim N(0, \sigma^2)$, $\sigma^2 = E(\epsilon^2) - E(\epsilon)^2 = E(\epsilon^2)$.
  Hence it can be estimated by:
  $$\widehat{\sigma^2} = s^2 = \frac{\sum_{i=1}^n e_i^2}{n-2}$$ → degree of freedom
  where $\hat{\sigma}$ is called residue standard error.
    - Log Likelihood Function:
      $$L(\sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$
      The maximum likelihood estimate (MLE) of σ^2 is given by:
      $$\hat{\sigma}^2_{MLE} = \frac{n-2}{n}s^2$$
- Fitted Regression Function: $\bar{Y} = \widehat{\beta_0} + \widehat{\beta_1}\bar{X}$
    - Fitted values: $\widehat{y_i} = \widehat{\beta_0} + \widehat{\beta_1} x_i$
    - Residues: $e_i = y_i - \widehat{y_i}$
- Sum of Squares
    - Total sum of squares (SST) estimates the variance of Y:
      $$\text{SST} = \sum_{i=1}^n (y_i - \bar{Y})^2$$
    - Regression sum of squares (SSR) is the variation of Y explained by X:
      $$\text{SSR} = \sum_{i=1}^n (\widehat{y_i} - \bar{Y})^2$$
    - Residue sum of squares (SSE) is the variation of Y caused by random errors:
      $$\text{SSE} = \sum_{i=1}^n (y_i - \widehat{y_i})^2$$
    - $\text{SST} \equiv \text{SSR} + \text{SSE}$

Source	d.f.	SS	MS	F
Regression	1	SSR	MSR	MSR/MSE
Error	$n-2$	SSE	MSE	
Total	$n-1$	SST		

- Coefficient of Determination: $R^2 = \frac{\text{SSR}}{\text{SST}}$
    - R^2 is the proportion of the variation in Y explained by X. It measures the strength of correlation between Y and X.
    - $R^2 = \text{corr}(Y, X)^2 = \text{corr}(Y, \hat{Y})^2$
    - Adjusted R^2: A less biased estimate.
      $$R_a^2 = \frac{n-1}{n-p-1}R^2 - \frac{p}{n-p-1}$$

1.3. Theories of Normal Distribution

- Normal Distribution: $f(y|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}$
    - Denotation: $N(\mu, \sigma^2)$
    - If $\mu = 0$, $\sigma^2 = 1$, it is called standard normal distribution.
- Multivariate Normal Distribution:
  $$f(\boldsymbol{z}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi|\Sigma|)^{\frac{q}{2}}} e^{-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{z}-\boldsymbol{\mu})}$$
    - Denotation: $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
    - If $\boldsymbol{Z}$ is a multivariate normal vector, then for any constant matrix $\boldsymbol{B}$, $\boldsymbol{BZ} \sim N(B\boldsymbol{\mu}, B\boldsymbol{\Sigma}B^{\mathrm{T}})$.
    - For any constant vector $\boldsymbol{c}$, the linear combination $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{Z}$ has a univariate normal distribution $N(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{\mu}, \boldsymbol{c}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{c})$.
- χ^2-distribution: Suppose $\boldsymbol{Z} = (Z_1, \ldots, Z_m)^T$ is a multivariate normal vector with $N(0, \boldsymbol{I})$. The distribution of $\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{Z} = \sum_{i=1}^m Z_i^2$ is called the χ^2-distribution with d.f. m and is denoted by χ^2_m.

- t-distribution: Suppose $Z \sim N(0,1)$, $U \sim \chi^2_m$, Z and U are independent. The distribution of $\frac{Z}{\sqrt{U/m}}$ is called the t-distribution with d.f. m and is denoted by t_m.
- F-distribution: Suppose $U \sim \chi^2_m$, $V \sim \chi^2_n$, U and V are independent. The distribution of $\frac{U/m}{V/n}$ is called the F-distribution of d.f. m and n, denoted by $F_{m,n}$.

1.4. Properties of Least Square Estimation

- Unbiasedness: $E(\widehat{\beta_1}) = \beta_1$, $E(\widehat{\beta_2}) = \beta_2$, $E(s^2) = \sigma^2$
- Properties of fitted parameters:
  - $\widehat{\beta_0} \sim N\left(\beta_0, \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right]\sigma^2\right)$
  - $\widehat{\beta_1} \sim N\left(\beta_1, \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right)$
  - $(n-2)\frac{s^2}{\sigma^2} \sim \chi^2_{n-2}$
- Properties of fitted value, $\hat{Y}$:
  - $E\hat{Y} = EY$
  - $\text{Var}(\hat{Y}) = \text{Var}(\widehat{\beta_0}) + 2X\text{Cov}(\widehat{\beta_0}, \widehat{\beta_1}) + X^2\text{Var}(\widehat{\beta_1})$

    $$= \left[\frac{1}{n} + \frac{(X - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right]\sigma^2$$
  - $E(Y - \hat{Y})^2 = \text{Var}(\hat{Y}) - \text{Var}(\epsilon)$

    $$= \left[1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right]\sigma^2$$

1.5. Statistical Inference for Simple Regression Model

- Significant Test: To test whether or not there is a linear regression relationship between the response variable and the covariate.
  - Null hypothesis: There is no such relationship.
  - Test statistics: $F = \frac{MSR}{MSE}$
  - Under the null hypothesis, $F \sim F_{1,n-2}$.
  - If $F > f_{1,n-2}(\alpha)$, the null hypothesis is rejected at level α. Otherwise, it is not rejected.
- Two-sided Significant Test for β_1:
  - Null hypothesis: $H_0: \beta_1 = 0$
  - Alternative hypothesis: $H_1: \beta \neq 0$
  - Test statistics: $T_{0\beta_1} = \frac{\widehat{\beta_1}}{s(\widehat{\beta_1})}$. Under H_0, $T_{0\beta 1} \sim t_{n-2}$.

    Here, $s^2(\beta_1) = \frac{s^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$.
  - p-value: $p = P(|t_{n-2}| > |T_{0\beta_1}|)$
  - Decision rule (with significance level α):
    1) If $|T_{0\beta_1}| > t_{n-2}\left(\frac{\alpha}{2}\right)$, reject H_0; otherwise, do not reject H_0.
    2) If $p < \alpha$, reject H_0; otherwise, do not reject H_0.
  - $100(1-\alpha)\%$ confidence interval for β_1:

    $$\left[\widehat{\beta_1} - t_{n-2}\left(\frac{\alpha}{2}\right)s(\widehat{\beta_1}), \widehat{\beta_1} + t_{n-2}\left(\frac{\alpha}{2}\right)s(\widehat{\beta_1})\right]$$
- One-sided Significant Test for β_1:
  - Hypotheses:
    1) $H_0: \beta_1 \leq 0$ vs $H_1: \beta_1 > 0$;
    2) $H_0: \beta_1 \geq 0$ vs $H_1: \beta_1 < 0$.

    Note that the equal sign must with H_0.
  - Decision rule:
    1) If $T_{0\beta 1} > t_{n-2}(\alpha)$, reject H_0. Otherwise, do not reject H_0.
    2) If $T_{0\beta 1} < -t_{n-2}(\alpha)$, reject H_0. Otherwise, do not reject H_0.
  - p-value:
    1) $p = P(t_{n-2} > T_{0\beta 1})$
    2) $p = P(t_{n-2} < T_{0\beta 1})$
  - $100(1-\alpha)\%$ lower confidence bound for β_1:

    $$\beta_1 \geq \widehat{\beta_1} - t_{n-2}(\alpha)s(\widehat{\beta_1})$$

    This corresponds to the one-sided test (1).

    $100(1-\alpha)\%$ lower confidence bound for β_1:

    $$\beta_1 \leq \widehat{\beta_1} + t_{n-2}(\alpha)s(\widehat{\beta_1})$$

    This corresponds to the one-sided test (2).
- Significant Test for β_0:
  - Test statistics: $T_{\beta_0} = \frac{\widehat{\beta_0} - \beta_0}{s(\widehat{\beta_0})}$. Under H_0, $T_{\beta_0} \sim t_{n-2}$.

    Here, $s^2(\widehat{\beta_0}) = s^2\left[\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right]$.
- Confidence Interval for $E(Y)$:

  $$\widehat{y_h} \pm t_{n-2}\left(\frac{\alpha}{2}\right)\sqrt{s^2\left(\frac{1}{n} + \frac{(x_h - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right)}$$
- Prediction Interval for Y_{new}:

  $$\widehat{y_h} \pm t_{n-2}\left(\frac{\alpha}{2}\right)\sqrt{s^2\left(1 + \frac{1}{n} + \frac{(x_h - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2}\right)}$$

1.6. *R* Commands and Interpretation of Results

Example – Simple Linear Regression Model

	impre	adver
1	32.1	50.1
2	99.6	74.1
3	11.7	19.3
4	21.9	22.9
5	60.8	82.4
6	78.6	40.1
7	92.4	185.9
8	50.7	26.9
9	21.4	20.4
10	40.1	166.2
11	40.8	27.0
12	10.4	45.6
13	88.9	154.9
14	12.0	5.0
15	29.2	49.7
16	38.0	26.9
17	10.0	5.7
18	12.3	7.6
19	23.4	9.2
20	71.1	32.4
21	4.4	6.1

Input the data into a data frame:

```
q1.dat=read.table(PATH_TO_FILE,header=TRUE)
```

Make the variables in the data frame available in the R console:

```
y=q1.dat$impre; x=q1.dat$adver
```

Plot `impre` on the vertical axis and `adver` on the horizontal axis:

```
plot(x, y, xlab=”Expenditure”, ylab=”Impression”)
```

Fit a simple linear regression to the data:

```
q1.fit = lm(y~x); summary(q1.fit)
```

Plot the estimated regression function:

```
abline(q1.fit,col=”blue”)
```

Plot the residue against the fitted values:

```
r = q1.fit$resid
fitted = q1.fit$fitted
plot(fitted, r, xlab=”Fitted values, ylab=”Residuals”)
```

Results:

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 22.16269    7.08948   3.126  0.00556 **
x            0.36317    0.09712   3.739  0.00139 **
```

Fitted regression function: $\hat{y} = 22.1627 + 0.36317x$

Predict a value: `predict(q1.fit, list(x=50))`

- `read.csv`: input data from Excel file with `.csv` format.
- `read.table`: input data from text file with `.txt` format.
- `plot(x, y)`: plot `y` against `x`.
- `object = lm(y~x)`: to call the `lm` to estimate the model and store the results in `object`.
- `summary(object)`: list the estimation results.
- `abline(object)`: draw the fitted regression line over the scatter plot of `y` againt `x`, it must be preceded by `plot(x,y)`.
- `fitted = object$fitted`: extract the fitted $\hat{y}$ and store them in `fitted`
- `r=residuals(object)`: extract the residuals and store them in `r`.
- Predicting at new covariate values:

```
predict(object, newdata,
        interval = c("none", "confidence", "prediction"),
        level = 0.95)
```

Interpretation:

- `Estimate`: (Intercept) — $\hat{\beta}_0$, X — $\hat{\beta}_1$.
- `Std. Error`: $sd(\hat{\beta}_0)$, $sd(\hat{\beta}_1)$.
- `t value`: $\hat{\beta}_0/sd(\hat{\beta}_0)$, $\hat{\beta}_1/sd(\hat{\beta}_1)$.
- `Pr(>|t|)`: p-values for two-sided tests on β_0 and β_1.
- `Residual standard error`: $\hat{\sigma}$.
- `Multiple R-squared`: R^2.
- `Adjusted R-squared`: R_a^2.
- `F-statistic`: MSR/MSE.
- `p-value`: p-value of the significant F test.

2. Multiple Linear Regression

2.1. Multiple LSM and its Estimation

- Multiple Linear Regression Model:

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \epsilon$$

  - Multiple Linear Regression Function:

$$EY = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p$$

    EY can also be written as $E(Y|X)$.

  - $\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$

$$X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1,p} \\ 1 & x_{21} & \cdots & x_{2,p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{n,p} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \boldsymbol{\epsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

- Least Square Estimation of $\boldsymbol{\beta}$:

$$\widehat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \boldsymbol{y}$$

- Least Square Estimation of σ^2:

$$\sigma^2 = \frac{\|\boldsymbol{y} - X(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\boldsymbol{y}\|^2}{n-p-1}$$

- Properties of Hat Matrix $H = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}$:
  - $HX = X$
  - $H^2 = H$
  - $(I-H)^2 = I - H$
  - Residue vector: $\boldsymbol{e} = \boldsymbol{y} - \widehat{\boldsymbol{y}} = (I-H)\boldsymbol{y}$
- Sum of Squares
  - SST: Let $\boldsymbol{y_t} = \begin{pmatrix} y_1 - \bar{y} \\ \dots \\ y_n - \bar{y} \end{pmatrix} = \boldsymbol{y} - \boldsymbol{1}\bar{y} \equiv H_t\boldsymbol{y}$.

$$\mathrm{SST} = \boldsymbol{y_t}^{\mathrm{T}}\boldsymbol{y_t} = \boldsymbol{y}^{\mathrm{T}}H_t^{\mathrm{T}}H_t\boldsymbol{y}$$

  - SSR:

$$\mathrm{SSR} = \boldsymbol{y_r}^{\mathrm{T}}\boldsymbol{y_r} = \boldsymbol{y}^{\mathrm{T}}H_r^{\mathrm{T}}H_r\boldsymbol{y}$$

  - SSE:

$$\mathrm{SSE} = \boldsymbol{e}^T\boldsymbol{e} = \boldsymbol{y}^{\mathrm{T}}H_e^{\mathrm{T}}H_e\boldsymbol{y}$$

  - $\mathrm{SST} = \mathrm{SSR} + \mathrm{SSE}$
  - d.f. of SST, SSR, SSE: $n-1$, p, $n-p-1$
- Distributions of Sum of Squares:
  - $\frac{\mathrm{SSE}}{\sigma^2} \sim \chi^2_{n-p-1}$
  - $\frac{\mathrm{SSR}}{\sigma^2} \sim \chi^2_p$

Source of Variation	SS	df	MS	F-statistic
Regression	SSR	p	$MSR = \frac{SSR}{p}$	MSR/MSE
Error	SSE	$n-p-1$	$MSE = \frac{SSE}{n-p-1}$	
Total	SST	$n-1$		

- Coefficient of Multiple Determination:

$$R^2 = \frac{\sum_{i=1}^n (\widehat{y_i} - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\mathrm{SSR}}{\mathrm{SST}}$$

  - $R^2 = \mathrm{corr}(\boldsymbol{y}, \widehat{\boldsymbol{y}})^2$
  - Adjusted R^2: $R_a^2 = R^2 - \frac{p}{n-p-1}(1-R^2)$

2.2. Inference on Least Square Estimation

- Properties of Least Square Estimation:
  - $\widehat{\boldsymbol{\beta}} \sim N(E(\widehat{\boldsymbol{\beta}}), Var(\widehat{\boldsymbol{\beta}}))$

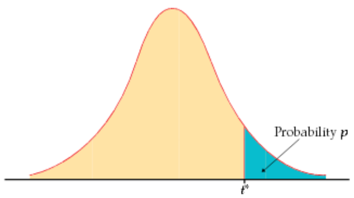
    $E(\widehat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$

    $\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \sigma^2(X^{\mathrm{T}}X)^{-1}$
  - $(n-p-1)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p-1}$
  - $\frac{MSR}{MSE} \sim F_{p,n-p-1}$
- Significance Test: Does the regression function have a significant effect on the response variable?
  - $H_0: \beta_1 = \cdots = \beta_p = 0$
  - H_1: At least one of $\beta_1, \dots, \beta_p \neq 0$
  - Test statistics: $F = \frac{MSR}{MSE}$. Under H_0, $F \sim F_{p,n-p-1}$.
  - Wald test statistics: $W = \widehat{\boldsymbol{\beta}}^{\mathrm{T}}\widehat{\boldsymbol{\Sigma}}_{\hat{\beta}}^{-1}\widehat{\boldsymbol{\beta}}$. $F = \frac{W}{p}$.
  - Decision rule: If $f > f_{p,n-p-1}(\alpha)$, reject H_0; otherwise, do not reject H_0.
- Individual t-test:
  - Hypothesis: $H_0: \beta_j = 0$ vs $H_1: \beta_j \neq 0$.
  - Test statistics: $t = \frac{\hat{\beta}_j}{\hat{\sigma}_j}$. Under H_0, $t \sim t_{n-p-1}$.
  - p-value: $p = 2P(t_{n-p-1} \geq t)$
  - Decision rule: If $|t| \geq t_{n-p-1}\left(\frac{\alpha}{2}\right)$, reject H_0; otherwise, do not reject H_0.
  - Confidence interval:

$$[\hat{\beta}_j - \hat{\sigma}_j t_{n-p-1}\left(\frac{\alpha}{2}\right), \hat{\beta}_j + \hat{\sigma}_j t_{n-p-1}\left(\frac{\alpha}{2}\right)]$$

t-distribution table

Areas in the upper tail are given along the top of the table. Critical t* values are given in the table.



df	0.1	0.05	0.025	0.02	0.01	0.005
1	3.078	6.314	12.706	15.895	31.821	63.657
2	1.886	2.920	4.303	4.849	6.965	9.925
3	1.638	2.353	3.182	3.482	4.541	5.841
4	1.533	2.132	2.776	2.999	3.747	4.604
5	1.476	2.015	2.571	2.757	3.365	4.032
6	1.440	1.943	2.447	2.612	3.143	3.707
7	1.415	1.895	2.365	2.517	2.998	3.499
8	1.397	1.860	2.306	2.449	2.896	3.355
9	1.383	1.833	2.262	2.398	2.821	3.250
10	1.372	1.812	2.228	2.359	2.764	3.169
11	1.363	1.796	2.201	2.328	2.718	3.106
12	1.356	1.782	2.179	2.303	2.681	3.055
13	1.350	1.771	2.160	2.282	2.650	3.012
14	1.345	1.761	2.145	2.264	2.624	2.977
15	1.341	1.753	2.131	2.249	2.602	2.947
16	1.337	1.746	2.120	2.235	2.583	2.921
17	1.333	1.740	2.110	2.224	2.567	2.898
18	1.330	1.734	2.101	2.214	2.552	2.878
19	1.328	1.729	2.093	2.205	2.539	2.861
20	1.325	1.725	2.086	2.197	2.528	2.845
21	1.323	1.721	2.080	2.189	2.518	2.831
22	1.321	1.717	2.074	2.183	2.508	2.819
23	1.319	1.714	2.069	2.177	2.500	2.807
24	1.318	1.711	2.064	2.172	2.492	2.797
25	1.316	1.708	2.060	2.167	2.485	2.787
26	1.315	1.706	2.056	2.162	2.479	2.779
27	1.314	1.703	2.052	2.158	2.473	2.771
28	1.313	1.701	2.048	2.154	2.467	2.763
29	1.311	1.699	2.045	2.150	2.462	2.756
30	1.310	1.697	2.042	2.147	2.457	2.750
31	1.309	1.696	2.040	2.144	2.453	2.744
32	1.309	1.694	2.037	2.141	2.449	2.738
33	1.308	1.692	2.035	2.138	2.445	2.733
34	1.307	1.691	2.032	2.136	2.441	2.728
35	1.306	1.690	2.030	2.133	2.438	2.724
36	1.306	1.688	2.028	2.131	2.434	2.719
37	1.305	1.687	2.026	2.129	2.431	2.715
38	1.304	1.686	2.024	2.127	2.429	2.712
39	1.304	1.685	2.023	2.125	2.426	2.708
40	1.303	1.684	2.021	2.123	2.423	2.704
41	1.303	1.683	2.020	2.121	2.421	2.701
42	1.302	1.682	2.018	2.120	2.418	2.698
43	1.302	1.681	2.017	2.118	2.416	2.695
44	1.301	1.680	2.015	2.116	2.414	2.692
45	1.301	1.679	2.014	2.115	2.412	2.690
46	1.300	1.679	2.013	2.114	2.410	2.687
47	1.300	1.678	2.012	2.112	2.408	2.685
48	1.299	1.677	2.011	2.111	2.407	2.682
49	1.299	1.677	2.010	2.110	2.405	2.680
50	1.299	1.676	2.009	2.109	2.403	2.678

df	0.1	0.05	0.025	0.02	0.01	0.005
51	1.298	1.675	2.008	2.108	2.402	2.676
52	1.298	1.675	2.007	2.107	2.400	2.674
53	1.298	1.674	2.006	2.106	2.399	2.672
54	1.297	1.674	2.005	2.105	2.397	2.670
55	1.297	1.673	2.004	2.104	2.396	2.668
56	1.297	1.673	2.003	2.103	2.395	2.667
57	1.297	1.672	2.002	2.102	2.394	2.665
58	1.296	1.672	2.002	2.101	2.392	2.663
59	1.296	1.671	2.001	2.100	2.391	2.662
60	1.296	1.671	2.000	2.099	2.390	2.660
61	1.296	1.670	2.000	2.099	2.389	2.659
62	1.295	1.670	1.999	2.098	2.388	2.657
63	1.295	1.669	1.998	2.097	2.387	2.656
64	1.295	1.669	1.998	2.096	2.386	2.655
65	1.295	1.669	1.997	2.096	2.385	2.654
66	1.295	1.668	1.997	2.095	2.384	2.652
67	1.294	1.668	1.996	2.095	2.383	2.651
68	1.294	1.668	1.995	2.094	2.382	2.650
69	1.294	1.667	1.995	2.093	2.382	2.649
70	1.294	1.667	1.994	2.093	2.381	2.648
71	1.294	1.667	1.994	2.092	2.380	2.647
72	1.293	1.666	1.993	2.092	2.379	2.646
73	1.293	1.666	1.993	2.091	2.379	2.645
74	1.293	1.666	1.993	2.091	2.378	2.644
75	1.293	1.665	1.992	2.090	2.377	2.643
76	1.293	1.665	1.992	2.090	2.376	2.642
77	1.293	1.665	1.991	2.089	2.376	2.641
78	1.292	1.665	1.991	2.089	2.375	2.640
79	1.292	1.664	1.990	2.088	2.374	2.640
80	1.292	1.664	1.990	2.088	2.374	2.639
81	1.292	1.664	1.990	2.087	2.373	2.638
82	1.292	1.664	1.989	2.087	2.373	2.637
83	1.292	1.663	1.989	2.087	2.372	2.636
84	1.292	1.663	1.989	2.086	2.372	2.636
85	1.292	1.663	1.988	2.086	2.371	2.635
86	1.291	1.663	1.988	2.085	2.370	2.634
87	1.291	1.663	1.988	2.085	2.370	2.634
88	1.291	1.662	1.987	2.085	2.369	2.633
89	1.291	1.662	1.987	2.084	2.369	2.632
90	1.291	1.662	1.987	2.084	2.368	2.632
91	1.291	1.662	1.986	2.084	2.368	2.631
92	1.291	1.662	1.986	2.083	2.368	2.630
93	1.291	1.661	1.986	2.083	2.367	2.630
94	1.291	1.661	1.986	2.083	2.367	2.629
95	1.291	1.661	1.985	2.082	2.366	2.629
96	1.290	1.661	1.985	2.082	2.366	2.628
97	1.290	1.661	1.985	2.082	2.365	2.627
98	1.290	1.661	1.984	2.081	2.365	2.627
99	1.290	1.660	1.984	2.081	2.365	2.626
100	1.290	1.660	1.984	2.081	2.364	2.626