

# Sujet numéro 1

Le sujet comporte trois exercices indépendants. Il est demandé à l'étudiant de ne traiter que deux exercices parmi les trois. La rédaction de trois exercices entraîne automatiquement un rejet de correction. Les calculatrices sont autorisées ainsi que les notes de cours, les tds et leurs corrigés. Des tables de Student, du  $\chi^2$  et de Fisher sont données en annexe du sujet.

## Exercice 1. Taille et masse des gousses de glycine blanche.

Un jardinier a ramassé dans son jardin des gousses de glycine blanche. Il cherche à savoir s'il est possible de déterminer de manière fiable la masse des gousses de glycine blanche à l'aide de leur taille ce qui lui éviterait d'avoir à les peser s'il les mesure. La masse des gousses est identifiée par  $Y$  et la taille des gousses par  $X$ .

Masse	28,6	20,6	29,2	32,0	24,5	29,0	28,9	18,2	7,9	15,5
Taille	19,1	14,8	19,7	21,1	19,4	19,5	18,9	14,6	10,2	14,6
Masse	22,6	35,5	32,5	28,7	26,0	13,5	16,4	12,5	26,2	22,6
Taille	16,4	21,1	20,7	18,7	17,6	13,2	14,0	12,0	18,3	17,8
Masse	9,7	21,8	17,2	25,2	12,0	6,3	7,0	20,4	18,0	21,1
Taille	10,7	16,5	14,5	17,5	12,2	8,6	9,1	17,0	15,3	15,8

- Le jardinier envisage d'utiliser le modèle linéaire simple pour modéliser le lien entre la masse et la taille des gousses. Spécifier ce modèle et bien identifier chacune des composantes du modèle dans le contexte de ce problème.
- Déterminer l'équation de la droite de régression.
- Donner une estimation de  $s(\beta_1)$ .
- Tester, au seuil de 5%, l'hypothèse nulle  $\mathcal{H}_0$  suivante avec un test approprié (vous donnerez le nom de ce test) :
 
$$\mathcal{H}_0 : \beta_1 = 0$$
 contre
 
$$\mathcal{H}_1 : \beta_1 \neq 0.$$
- Donner la valeur de la variation qui est expliquée par la droite de régression et la variation qui est inexpliquée par la même droite.
- Donner le pourcentage de variation qui est expliqué par la droite de régression.
- Le modèle envisagé par le jardinier vous semble-t-il pertinent ? Vous détaillerez votre réponse en utilisant certains des résultats des questions précédentes.
- Calculer une estimation de la masse des gousses pour les valeurs de tailles suivantes :  $X_1 = 14,8$ ;  $X_2 = 16,4$ ;  $X_3 = 12,0$ .
- Pour quelle quantité  $X_i$  l'estimation de la masse des gousses serait-elle la plus précise parmi les trois valeurs précédentes ?
- Entre quelles valeurs peut se situer la vraie valeur de la masse pour les gousses dont la masse a été déterminée à la question précédente ? Utiliser un niveau de confiance de 99%.

11. Quelle est la marge d'erreur dans l'estimation effectuée à la question précédente?

```
> masse1<-c(28.6,20.6,29.2,32.0,24.5,29.0,28.9,18.2,7.9,15.5,22.6,
+ 35.5,32.5,28.7,26.0,13.5,16.4,12.5,26.2,22.6,9.7,21.8,17.2,25.2,
+ 12.0,6.3,7.0,20.4,18.0,21.1)
> taille1<-c(19.1,14.8,19.7,21.1,19.4,19.5,18.9,14.6,10.2,14.6,16.4,
+ 21.1,20.7,18.7,17.6,13.2,14.0,12.0,18.3,17.8,10.7,16.5,14.5,17.5,
+ 12.2,8.6,9.1,17.0,15.3,15.8)
```

```
> model<-lm(masse1~taille1)
> residus<-residuals(model)
> shapiro.test(residus)
      Shapiro-Wilk normality test
```

```
data: residus
W = 0.9606, p-value = 0.3217
```

```
> summary(model)
Call:
lm(formula = masse1 ~ taille1)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-4.0958 -0.5037  0.2073  0.9721  3.1402
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -14.35790    1.36014  -10.56 2.88e-11 ***
taille1       2.21411    0.08323   26.60 < 2e-16 ***
---

```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.593 on 28 degrees of freedom
Multiple R-squared: 0.9619, Adjusted R-squared: 0.9606
F-statistic: 707.6 on 1 and 28 DF, p-value: < 2.2e-16
```

```
> anova(model)
Analysis of Variance Table
```

```
Response: masse1
      Df Sum Sq Mean Sq F value    Pr(>F)
taille1  1 1794.97 1794.97  707.61 < 2.2e-16 ***
Residuals 28   71.03    2.54
---

```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> sum(taille1)
[1] 478.9
```

```
> sum(taille1^2)
[1] 8010.99
```

### Exercice 2. Nombre de graines et glycine.

Les graines des glycines sont contenues dans des gousses. Il y a généralement plusieurs graines dans une même gousse. Nous nous proposons de déterminer si le nombre de graines dans une gousse de glycine dépend de l'espèce de glycine considérée. L'expérience a ici été réalisée avec des glycines blanches et des glycines violettes. Nous regroupons chacun des échantillons obtenus en quatre catégories :

- $A$  : Inférieur ou égal à trois graines
- $B$  : Quatre graines
- $C$  : Cinq graines
- $D$  : Supérieur ou égal à six graines

Les résultats obtenus figurent dans le tableau suivant :

Espèce	A	B	C	D
Glycine blanche	7	19	11	10
Glycine violette	4	22	16	6

Les espèces sont-elles significativement différentes ?

Pour répondre à la question, vous effectuerez un test dont vous donnerez le nom, puis vous énoncerez les deux hypothèses associées à ce test ainsi que la valeur de la statistique de ce test. Enfin, il manque deux valeurs dans la sortie de R, retrouvez ces valeurs ainsi que la valeur critique du test.

```
> plantes<-matrix(c(7,19,11,10,4,22,16,6),byrow=T,nrow=2,
+ dimnames=list(c("Blanche","Violette"),c("A","B","C","D")))
> plantes
```

```
      A  B  C  D
Blanche  7 19 11 10
Violette  4 22 16  6
```

```
> chisq.test(plantes,correct=FALSE)
```

Pearson's Chi-squared test

```
data:  plantes
X-squared = 2.9534, df = 3, p-value = 0.3989
```

```
> chisq.test(plantes,correct=FALSE)$expected
      A      B      C      D
Blanche  ? 20.28421 13.35789 7.91579
Violette  ? 20.71579 13.64211 8.08421
```

**Exercice 3. Les graines.**

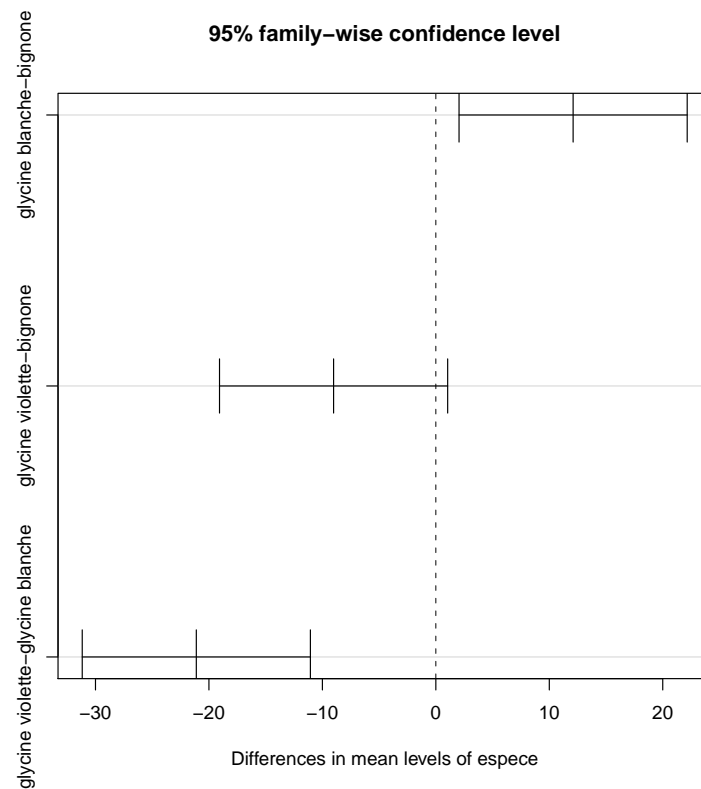
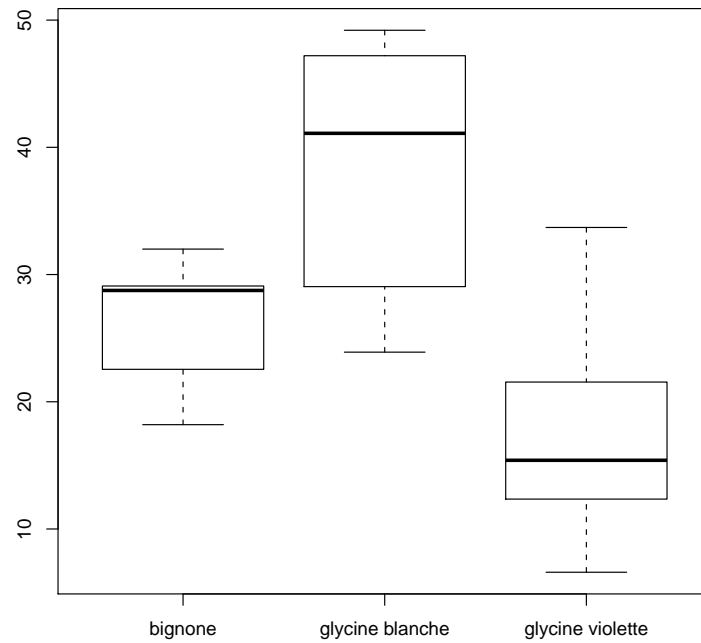
Nous cherchons à déterminer si les masses moyennes des graines de glycine blanche, des graines de glycine violette et des graines de bignone sont égales.

Nous disposons de 24 graines de ces trois plantes : huit graines de glycine blanche, huit graines de glycine violette et huit graines de bignone.

Le tableau suivant donne la masse, en grammes, de chacune des graines :

Espèce	Glycine blanche	Glycine violette	Bignone
$x_{i1}$	28.6	40.0	10.9
$x_{i2}$	20.6	49.2	6.6
$x_{i3}$	29.2	46.0	22.5
$x_{i4}$	32.0	26.4	33.7
$x_{i5}$	24.5	42.2	20.6
$x_{i6}$	29.0	48.4	16.6
$x_{i7}$	28.9	23.9	14.2
$x_{i8}$	18.2	31.7	13.8

1. Quelle est la nature des variables présentes dans le jeu de données ?
2. Quel est le modèle statistique qui a été utilisé pour analyser ces données ?
3. Les conditions d'utilisation de ce modèle sont-elles remplies ?
4. Donner le tableau de l'analyse de la variance.
5. Donner les valeurs des coefficients du modèle.
6. Quelle est la valeur estimée de la variance résiduelle ?
7. Au risque de 5%, pouvez-vous considérer que l'espèce possède une influence sur la masse des graines ?
8. Pouvez-vous séparer les espèces en groupes ne présentant pas de différence significative au seuil de 5% ? Si oui, expliquer comment vous procédez.
9. Dans le cas où vous avez répondu dans l'affirmative à la question précédente, faire cette répartition en groupes homogènes en indiquant les espèces et les moyennes correspondantes.



```
> options(contrasts=c("contr.sum", "contr.poly"))
> masse1<-c(28.6,20.6,29.2,32.0,24.5,29.0,28.9,18.2,40.0,49.2,46.0,26.4,
+ 42.2,48.4,23.9,31.7,10.9,6.6,22.5,33.7,20.6,16.6,14.2,13.8)
```

```
> espece<-rep(c("bignone","glycine blanche","glycine violette"),c(8,8,8))
> espece<-factor(espece)
> espece
> moys<-tapply(masse1,espece,mean)
> moys
      bignone  glycine blanche  glycine violette
      26.3750         38.4750         17.3625

> sds<-tapply(masse1,espece,sd)
> sds
      bignone  glycine blanche  glycine violette
      4.803793         9.931875         8.320875

> model1<-lm(masse1~espece)
> model1
Call:
lm(formula = masse1 ~ espece)

Coefficients:
(Intercept)      espece1      espece2
      27.404         -1.029         11.071

> residus<-residuals(model1)
> shapiro.test(residus)

      Shapiro-Wilk normality test

data:  residus
W = 0.9846, p-value = 0.963

> bartlett.test(residus~espece)

      Bartlett test of homogeneity of variances

data:  residus by espece
Bartlett's K-squared = 3.2413, df = 2, p-value = 0.1978

> summary(model1)

Call:
lm(formula = masse1 ~ espece)

Residuals:
      Min       1Q   Median       3Q      Max
-14.575  -5.947   1.875   4.078  16.338

Coefficients:
```

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept)  27.404      1.629  16.827 1.15e-13 ***
espece1      -1.029      2.303  -0.447   0.66
espece2      11.071      2.303   4.807 9.47e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 7.978 on 21 degrees of freedom
Multiple R-squared:  0.5733,    Adjusted R-squared:  0.5326
F-statistic: 14.11 on 2 and 21 DF,  p-value: 0.0001308

```

```

> anova(model1)
Analysis of Variance Table

```

```

Response: masse1
      Df Sum Sq Mean Sq F value    Pr(>F)
espece  2 1795.7   897.83  14.105 0.0001308 ***
Residuals 21 1336.7    63.65
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

> model2<-aov(masse1~espece)

```

```

> TukeyHSD(model2)
Tukey multiple comparisons of means
 95% family-wise confidence level

```

```

Fit: aov(formula = masse1 ~ espece)

```

```

$espece
              diff          lwr          upr          p adj
glycine blanche-bignone      12.1000    2.04518   22.154820 0.0166783
glycine violette-bignone     -9.0125  -19.06732    1.042320 0.0842656
glycine violette-glycine blanche -21.1125  -31.16732  -11.057680 0.0000862

```

## Table de la loi de Student

$\alpha$	0.900	0.950	0.975	0.990	0.995
1	3.0777	6.3138	12.7062	31.8205	63.6567
2	1.8856	2.9200	4.3027	6.9646	9.9248
3	1.6377	2.3534	3.1824	4.5407	5.8409
4	1.5332	2.1318	2.7764	3.7469	4.6041
5	1.4759	2.0150	2.5706	3.3649	4.0321
6	1.4398	1.9432	2.4469	3.1427	3.7074
7	1.4149	1.8946	2.3646	2.9980	3.4995
8	1.3968	1.8595	2.3060	2.8965	3.3554
9	1.3830	1.8331	2.2622	2.8214	3.2498
10	1.3722	1.8125	2.2281	2.7638	3.1693
11	1.3634	1.7959	2.2010	2.7181	3.1058
12	1.3562	1.7823	2.1788	2.6810	3.0545
13	1.3502	1.7709	2.1604	2.6503	3.0123
14	1.3450	1.7613	2.1448	2.6245	2.9768
15	1.3406	1.7531	2.1314	2.6025	2.9467
16	1.3368	1.7459	2.1199	2.5835	2.9208
17	1.3334	1.7396	2.1098	2.5669	2.8982
18	1.3304	1.7341	2.1009	2.5524	2.8784
19	1.3277	1.7291	2.0930	2.5395	2.8609
20	1.3253	1.7247	2.0860	2.5280	2.8453
21	1.3232	1.7207	2.0796	2.5176	2.8314
22	1.3212	1.7171	2.0739	2.5083	2.8188
23	1.3195	1.7139	2.0687	2.4999	2.8073
24	1.3178	1.7109	2.0639	2.4922	2.7969
25	1.3163	1.7081	2.0595	2.4851	2.7874
26	1.3150	1.7056	2.0555	2.4786	2.7787
27	1.3137	1.7033	2.0518	2.4727	2.7707
28	1.3125	1.7011	2.0484	2.4671	2.7633
29	1.3114	1.6991	2.0452	2.4620	2.7564
30	1.3104	1.6973	2.0423	2.4573	2.7500



## Table de la loi du $\chi^2$ , régions unilatérales

$\alpha$	0.01	0.05	0.10	0.90	0.95	0.99
1	0.0002	0.0039	0.0158	2.7055	3.8415	6.6349
2	0.0201	0.1026	0.2107	4.6052	5.9915	9.2103
3	0.1148	0.3518	0.5844	6.2514	7.8147	11.3449
4	0.2971	0.7107	1.0636	7.7794	9.4877	13.2767
5	0.5543	1.1455	1.6103	9.2364	11.0705	15.0863
6	0.8721	1.6354	2.2041	10.6446	12.5916	16.8119
7	1.2390	2.1673	2.8331	12.0170	14.0671	18.4753
8	1.6465	2.7326	3.4895	13.3616	15.5073	20.0902
9	2.0879	3.3251	4.1682	14.6837	16.9190	21.6660
10	2.5582	3.9403	4.8652	15.9872	18.3070	23.2093
11	3.0535	4.5748	5.5778	17.2750	19.6751	24.7250
12	3.5706	5.2260	6.3038	18.5493	21.0261	26.2170
13	4.1069	5.8919	7.0415	19.8119	22.3620	27.6882
14	4.6604	6.5706	7.7895	21.0641	23.6848	29.1412
15	5.2293	7.2609	8.5468	22.3071	24.9958	30.5779
16	5.8122	7.9616	9.3122	23.5418	26.2962	31.9999
17	6.4078	8.6718	10.0852	24.7690	27.5871	33.4087
18	7.0149	9.3905	10.8649	25.9894	28.8693	34.8053
19	7.6327	10.1170	11.6509	27.2036	30.1435	36.1909
20	8.2604	10.8508	12.4426	28.4120	31.4104	37.5662
21	8.8972	11.5913	13.2396	29.6151	32.6706	38.9322
22	9.5425	12.3380	14.0415	30.8133	33.9244	40.2894
23	10.1957	13.0905	14.8480	32.0069	35.1725	41.6384
24	10.8564	13.8484	15.6587	33.1962	36.4150	42.9798
25	11.5240	14.6114	16.4734	34.3816	37.6525	44.3141
26	12.1981	15.3792	17.2919	35.5632	38.8851	45.6417
27	12.8785	16.1514	18.1139	36.7412	40.1133	46.9629
28	13.5647	16.9279	18.9392	37.9159	41.3371	48.2782
29	14.2565	17.7084	19.7677	39.0875	42.5570	49.5879
30	14.9535	18.4927	20.5992	40.2560	43.7730	50.8922

## Table de la loi de Fisher-Snedecor, $\alpha = 5\%$

num	den 1	2	3	4	5	6	7	8	9	10
1	161.4476	18.5128	10.1280	7.7086	6.6079	5.9874	5.5914	5.3177	5.1174	4.9646
2	199.5000	19.0000	9.5521	6.9443	5.7861	5.1433	4.7374	4.4590	4.2565	4.1028
3	215.7073	19.1643	9.2766	6.5914	5.4095	4.7571	4.3468	4.0662	3.8625	3.7083
4	224.5832	19.2468	9.1172	6.3882	5.1922	4.5337	4.1203	3.8379	3.6331	3.4780
5	230.1619	19.2964	9.0135	6.2561	5.0503	4.3874	3.9715	3.6875	3.4817	3.3258
6	233.9860	19.3295	8.9406	6.1631	4.9503	4.2839	3.8660	3.5806	3.3738	3.2172
7	236.7684	19.3532	8.8867	6.0942	4.8759	4.2067	3.7870	3.5005	3.2927	3.1355
8	238.8827	19.3710	8.8452	6.0410	4.8183	4.1468	3.7257	3.4381	3.2296	3.0717
9	240.5433	19.3848	8.8123	5.9988	4.7725	4.0990	3.6767	3.3881	3.1789	3.0204
10	241.8817	19.3959	8.7855	5.9644	4.7351	4.0600	3.6365	3.3472	3.1373	2.9782
11	242.9835	19.4050	8.7633	5.9358	4.7040	4.0274	3.6030	3.3130	3.1025	2.9430
12	243.9060	19.4125	8.7446	5.9117	4.6777	3.9999	3.5747	3.2839	3.0729	2.9130
13	244.6898	19.4189	8.7287	5.8911	4.6552	3.9764	3.5503	3.2590	3.0475	2.8872
14	245.3640	19.4244	8.7149	5.8733	4.6358	3.9559	3.5292	3.2374	3.0255	2.8647
15	245.9499	19.4291	8.7029	5.8578	4.6188	3.9381	3.5107	3.2184	3.0061	2.8450
16	246.4639	19.4333	8.6923	5.8441	4.6038	3.9223	3.4944	3.2016	2.9890	2.8276
17	246.9184	19.4370	8.6829	5.8320	4.5904	3.9083	3.4799	3.1867	2.9737	2.8120
18	247.3232	19.4402	8.6745	5.8211	4.5785	3.8957	3.4669	3.1733	2.9600	2.7980
19	247.6861	19.4431	8.6670	5.8114	4.5678	3.8844	3.4551	3.1613	2.9477	2.7854
20	248.0131	19.4458	8.6602	5.8025	4.5581	3.8742	3.4445	3.1503	2.9365	2.7740
21	248.3094	19.4481	8.6540	5.7945	4.5493	3.8649	3.4349	3.1404	2.9263	2.7636
22	248.5791	19.4503	8.6484	5.7872	4.5413	3.8564	3.4260	3.1313	2.9169	2.7541
23	248.8256	19.4523	8.6432	5.7805	4.5339	3.8486	3.4179	3.1229	2.9084	2.7453
24	249.0518	19.4541	8.6385	5.7744	4.5272	3.8415	3.4105	3.1152	2.9005	2.7372
25	249.2601	19.4558	8.6341	5.7687	4.5209	3.8348	3.4036	3.1081	2.8932	2.7298
26	249.4525	19.4573	8.6301	5.7635	4.5151	3.8287	3.3972	3.1015	2.8864	2.7229
27	249.6309	19.4587	8.6263	5.7586	4.5097	3.8230	3.3913	3.0954	2.8801	2.7164
28	249.7966	19.4600	8.6229	5.7541	4.5047	3.8177	3.3858	3.0897	2.8743	2.7104
29	249.9510	19.4613	8.6196	5.7498	4.5001	3.8128	3.3806	3.0844	2.8688	2.7048
30	250.0951	19.4624	8.6166	5.7459	4.4957	3.8082	3.3758	3.0794	2.8637	2.6996

num	den 11	12	13	14	15	16	17	18	19	20
1	4.8443	4.7472	4.6672	4.6001	4.5431	4.4940	4.4513	4.4139	4.3807	4.3512
2	3.9823	3.8853	3.8056	3.7389	3.6823	3.6337	3.5915	3.5546	3.5219	3.4928
3	3.5874	3.4903	3.4105	3.3439	3.2874	3.2389	3.1968	3.1599	3.1274	3.0984
4	3.3567	3.2592	3.1791	3.1122	3.0556	3.0069	2.9647	2.9277	2.8951	2.8661
5	3.2039	3.1059	3.0254	2.9582	2.9013	2.8524	2.8100	2.7729	2.7401	2.7109
6	3.0946	2.9961	2.9153	2.8477	2.7905	2.7413	2.6987	2.6613	2.6283	2.5990
7	3.0123	2.9134	2.8321	2.7642	2.7066	2.6572	2.6143	2.5767	2.5435	2.5140
8	2.9480	2.8486	2.7669	2.6987	2.6408	2.5911	2.5480	2.5102	2.4768	2.4471
9	2.8962	2.7964	2.7144	2.6458	2.5876	2.5377	2.4943	2.4563	2.4227	2.3928
10	2.8536	2.7534	2.6710	2.6022	2.5437	2.4935	2.4499	2.4117	2.3779	2.3479
11	2.8179	2.7173	2.6347	2.5655	2.5068	2.4564	2.4126	2.3742	2.3402	2.3100
12	2.7876	2.6866	2.6037	2.5342	2.4753	2.4247	2.3807	2.3421	2.3080	2.2776
13	2.7614	2.6602	2.5769	2.5073	2.4481	2.3973	2.3531	2.3143	2.2800	2.2495
14	2.7386	2.6371	2.5536	2.4837	2.4244	2.3733	2.3290	2.2900	2.2556	2.2250
15	2.7186	2.6169	2.5331	2.4630	2.4034	2.3522	2.3077	2.2686	2.2341	2.2033
16	2.7009	2.5989	2.5149	2.4446	2.3849	2.3335	2.2888	2.2496	2.2149	2.1840
17	2.6851	2.5828	2.4987	2.4282	2.3683	2.3167	2.2719	2.2325	2.1977	2.1667
18	2.6709	2.5684	2.4841	2.4134	2.3533	2.3016	2.2567	2.2172	2.1823	2.1511
19	2.6581	2.5554	2.4709	2.4000	2.3398	2.2880	2.2429	2.2033	2.1683	2.1370
20	2.6464	2.5436	2.4589	2.3879	2.3275	2.2756	2.2304	2.1906	2.1555	2.1242
21	2.6358	2.5328	2.4479	2.3768	2.3163	2.2642	2.2189	2.1791	2.1438	2.1124
22	2.6261	2.5229	2.4379	2.3667	2.3060	2.2538	2.2084	2.1685	2.1331	2.1016
23	2.6172	2.5139	2.4287	2.3573	2.2966	2.2443	2.1987	2.1587	2.1233	2.0917
24	2.6090	2.5055	2.4202	2.3487	2.2878	2.2354	2.1898	2.1497	2.1141	2.0825
25	2.6014	2.4977	2.4123	2.3407	2.2797	2.2272	2.1815	2.1413	2.1057	2.0739
26	2.5943	2.4905	2.4050	2.3333	2.2722	2.2196	2.1738	2.1335	2.0978	2.0660
27	2.5877	2.4838	2.3982	2.3264	2.2652	2.2125	2.1666	2.1262	2.0905	2.0586
28	2.5816	2.4776	2.3918	2.3199	2.2587	2.2059	2.1599	2.1195	2.0836	2.0517
29	2.5759	2.4718	2.3859	2.3139	2.2525	2.1997	2.1536	2.1131	2.0772	2.0452
30	2.5705	2.4663	2.3803	2.3082	2.2468	2.1938	2.1477	2.1071	2.0712	2.0391