NTNU
Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for **TMA4275 Lifetime Analysis**

Academic contact during examination: Jacopo Paglia

Phone: 967 03 414

Examination date: Friday, June 7, 2019

Examination time (from–to): 09:00-13:00

Permitted examination support material: C: Approved Calculator. One yellow sheet (A4 with stamp) with your own formulae and notes.

Other information:
Tables of the standard normal distribution are enclosed at the end of the exam.

Language: English

Number of pages: 8

Number of pages enclosed: 3

Checked by:

Informasjon om trykking av eksamensoppgave
Originalen er:
1-sidig ☐ **2-sidig** ☒
sort/hvit ☐ **farger** ☒
skal ha flervalgskjema ☐

Date Signature

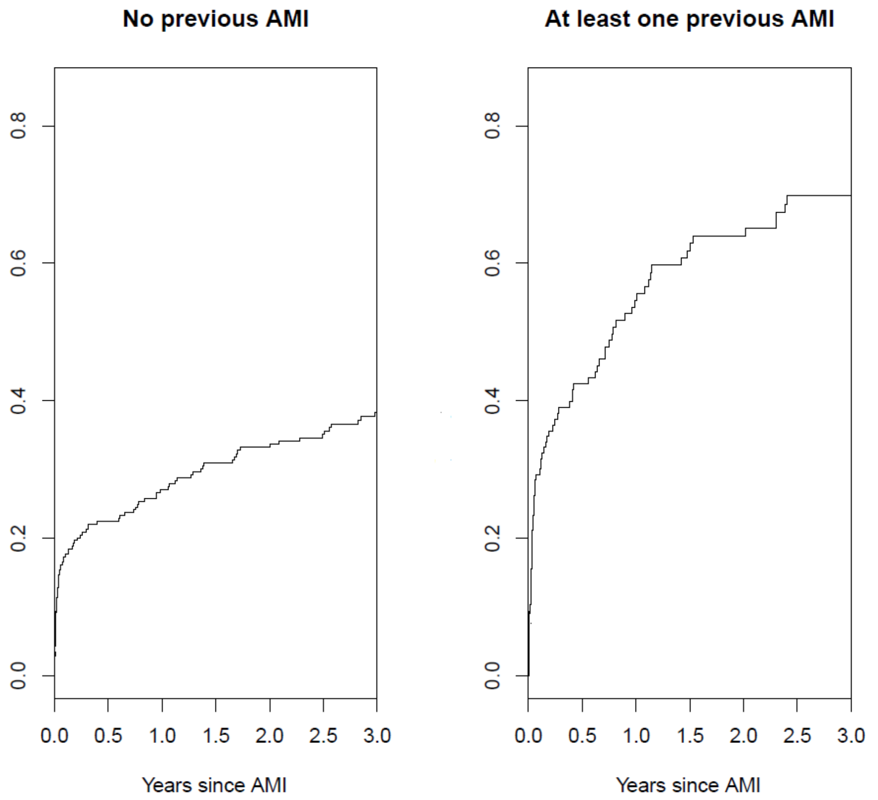

Figure 1: Nelson-Aalen plots for the two groups of AMI patients.

Problem 1

a) Suppose you are given the hazard function $z(t)$ of a lifetime T. Express (without proof) the following functions using only the function $z(t)$:

- the cumulative hazard function $Z(t)$,
- the density function $f(t)$,
- the reliability (survival) function $R(t)$.

Which of these functions is estimated by the Nelson-Aalen estimator?

Write down an expression for the Nelson-Aalen estimator based on a set of right censored lifetimes. Which are the standard conditions for using this estimator?

At a hospital one has studied the survival of patients admitted with an acute myocardial infarction (AMI). Figure 1 shows Nelson-Aalen estimates for patients with no previous AMI (left) and at least one previous AMI (right).

Use the plots to make rough sketches of the hazard functions $z(t)$ for the two groups of patients. Discuss what the sketches (and the Nelson-Aalen plots) tell you about the mortality of AMI patients.

Problem 2

a) What is the purpose of accelerated life testing (ALT)?

A commonly used model in ALT is the Arrhenius model, which is given by

$$\ln T = \beta_0 + \beta_1 \cdot \frac{11604.83}{s + 273.16} + \sigma U. \tag{1}$$

Explain what the different variables and parameters mean in this equation.

Briefly explain how the model will typically be used in a practical lifetime study.

For a particular product, a requirement has been set that at least 95% of the manufactured units should have a lifetime exceeding 30 000 hours at a temperature of 10^{o}C.

An experiment was performed with accelerated life testing, where a total of 165 units of the product were tested at temperatures 10^{o}C, 40^{o}C, 60^{o}C, 80^{o}C. The experiment was terminated after 5000 hours. Then 33 units had failed, while 132 still worked. The observations are given in Table 1 on page 4. The table gives observed time; censoring status (0 for right censoring at 5000 hours, 1 for observed lifetime); frequency (greater than 1 if several units had the same value for time, censoring status and temperature); test temperature in ^{o}C. A plot of the observations is given in Figure 2.

The analysis of the data was based on the model (1), where it was assumed that the lifetime T of a unit is lognormally distributed, while s is the test temperature measured in ^{o}C.

An output from MINITAB is provided below. You shall use results from this in the rest of the problem.

```
Accelerated Life Testing: Time versus Temp

Response Variable: Time
Frequency: Frequency

Censoring Information  Count
Uncensored value          33
Right censored value     132

Censoring value: Censor = 0

Estimation Method: Maximum Likelihood
```

```
Distribution:   Lognormal

Relationship with accelerating variable(s):   Arrhenius


Regression Table
```

Predictor	Coef	Standard Error	Z	P	95,0% Normal CI Lower	Upper
Intercept	-13,4693	2,88728	-4,67	0,000	-19,1283	-7,81034
Temp	0,627900	0,0828450	7,58	0,000	0,465527	0,790273
Scale	0,977823	0,132647			0,749532	1,27565

```
Log-Likelihood = -321,703
```

b) Write down the estimated model.

What is the distribution of U under the given assumption of lognormal distribution?

Find an estimate for $P(T > 30\ 000)$ at temperature 10^o C. Compare with the given reliability requirement (see above). Comment.

Find estimates of the expectation and median of the product's lifetime at 10^o C.

Also estimate the 5% quantile for the lifetime of the product at 10^o C, i.e. the time that 95% of the produced units will survive.

c) Figure 3 shows a probability plot for the data. Explain how the points of this plot are generated. You should do this by deriving the relevant formulas for the plotted points and the corresponding reference lines. (*Hint:* Start from the cumulative distribution function $F(t)$ for a general lognormal distribution.)

Why are the reference lines for the different temperatures parallel?

What can you read from this plot?

What is meant by standardized residuals for the given model? Explain how the residual plot of standardized residuals in Figure 4 is generated.

What do the plots in Figure 3 and 4 tell about the model's fit to the given failure data?

Row	Time	Censor	Frequency	Temp
1	5000	0	30	10
2	1298	1	1	40
3	1390	1	1	40
4	3187	1	1	40
5	3241	1	1	40
6	3261	1	1	40
7	3313	1	1	40
8	4501	1	1	40
9	4568	1	1	40
10	4841	1	1	40
11	4982	1	1	40
12	5000	0	90	40
13	581	1	1	60
14	925	1	1	60
15	1432	1	1	60
16	1586	1	1	60
17	2452	1	1	60
18	2734	1	1	60
19	2772	1	1	60
20	4106	1	1	60
21	4674	1	1	60
22	5000	0	11	60
23	283	1	1	80
24	361	1	1	80
25	515	1	1	80
26	638	1	1	80
27	854	1	1	80
28	1024	1	1	80
29	1030	1	1	80
30	1045	1	1	80
31	1767	1	1	80
32	1777	1	1	80
33	1856	1	1	80
34	1951	1	1	80
35	1964	1	1	80
36	2884	1	1	80
37	5000	0	1	80

Table 1: The observations from the ALT experiment.

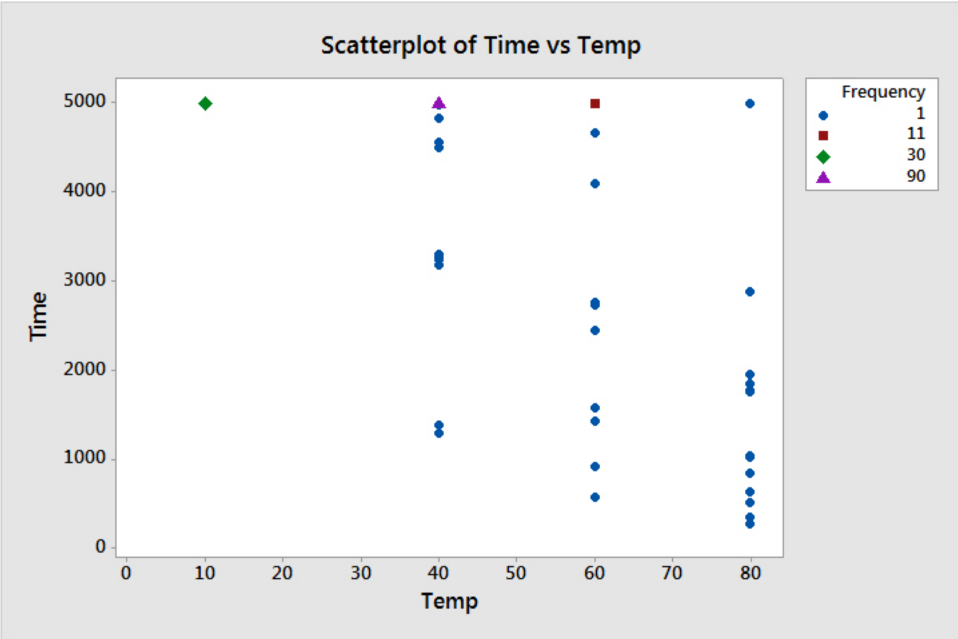

Figure 2: Scatterplot for the observed times.

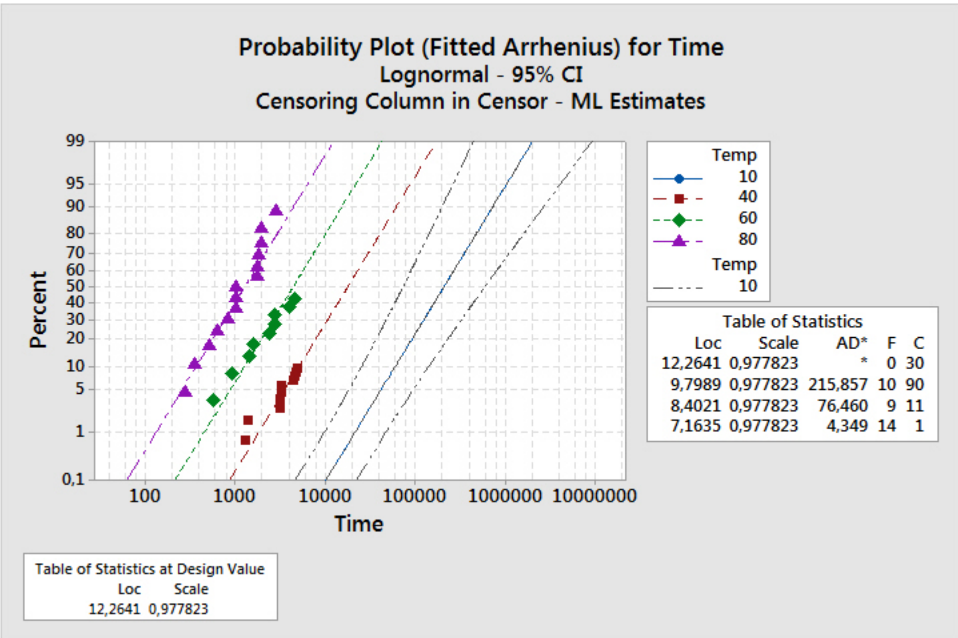

Figure 3: Probability plot for the observed times.

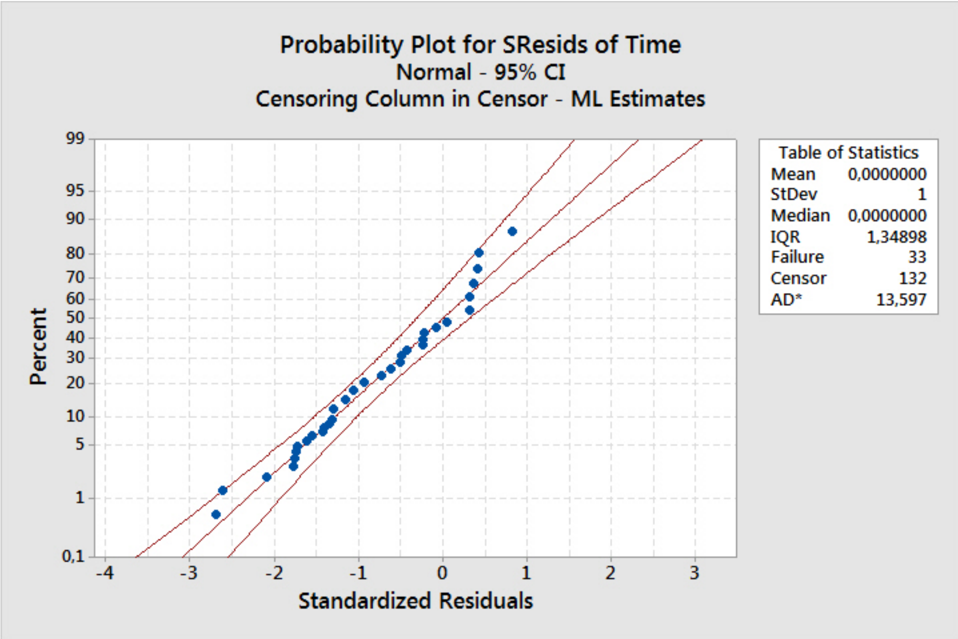

Figure 4: Probability plot for standardized residuals.

Problem 3

A software system is subject to failures at random times caused by errors present in the code. Let $N(t)$ be the cumulative number of failures experienced by time $t \geq 0$. We assume that each failure is caused by exactly one error in the code, and that this error is successfully removed from the software after each failure. Hence $N(t)$ also represents the cumulative number of errors detected and removed by time t.

A classical model for software reliability is the *time dependent error detection model of Goel and Okumoto* (the GO-model), where one assumes that $N(t)$ is an NHPP with intensity function

$$w(t) = \alpha\beta e^{-\beta t} \text{ for } t \geq 0, \tag{2}$$

where $\alpha > 0, \beta > 0$ are parameters.

In questions (a) and (b) are assumed that the error detection process $N(t)$ follows the GO-model (2).

a) Find the cumulative intensity function $W(t)$. Can you give an interpretation of $W(t)$ in the given situation?

Find the probability that the first error in the code is found before time s, where $s > 0$ is a given time.

What is the probability that the first error is found in the time interval from s to t, where $0 < s < t$?

What is the limit of $W(t)$ when $t \to \infty$? Explain why this limit provides a reasonable interpretation of the parameter α as the "initial number of errors" in the code.

b) Suppose that the error detection process has been run until time $s > 0$. The *conditional reliability function* $R(t|s)$ of the software at time s is defined to be the probability that the software operates without failures for at least time t beyond time s. Show that

$$R(t|s) = \exp\left(-e^{-\beta s} W(t)\right) \text{ for } t > 0,\ s > 0. \tag{3}$$

Use (3) to calculate an *optimal testing time* $s = s_0$ by the requirement on the testing time s, that the conditional reliability $R(t_0|s)$ should be at least equal to a given value r $(0 < r < 1)$, for a given time $t_0 > 0$. Express s_0 in terms of α, β, r and t_0. Discuss in particular the choice $t_0 = \infty$.

Recall the interpretation of the parameter α as the "initial number of errors" in the code. In the *Jelinski-Moranda model for software reliability* (the JM-model) the *true* initial number of errors, a, is an unknown parameter, which is hence a non-negative integer which we shall assume is at least 1.

The basic assumption of the JM-model is that the time T_i between failure number $i - 1$ and failure number i is exponentially distributed with hazard rate

$$\lambda_i = (a - i + 1)b \tag{4}$$

for $i = 1, 2, \ldots, a$, where $b > 0$ is the second parameter of the model. Failure number 0 will here correspond to the time when the software testing is started, i.e. time $t = 0$. It is further assumed that $T_1, T_2, \ldots, T_a$ are *stochastically independent.*

In the rest of the problem, the error detection process $N(t)$ is assumed to follow the JM model (4).

c) Explain why the time S_a of the ath failure can be interpreted as the time when all errors in the code are found.

Derive an exact expression for $E(S_a)$, i.e. the expected time when the software is error-free.

Then verify the approximation

$$E(S_a) \approx \frac{\ln a}{b}$$

when a is large. (*Hint:* A sum may be approximated by an integral).

Compare this approximate value to the optimal testing time s_0 considered in question (b) and comment.

d) A test procedure for a new computer program consists in running the program until a given number, m, of errors are found. Let the observed times for this be $s_1, s_2, \ldots, s_m$.

Write down the likelihood function $L(a, b)$ for the JM model based on these observations.

Briefly explain how you would proceed to find the maximum likelihood estimates for a and b.

Standard normalfordeling

$$\Phi(z) = P(Z \leq z)$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Standard normalfordeling

$$\Phi(z) = P(Z \leq z)$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.7	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Kritiske verdier i standard normalfordelingen

$$P(Z > z_\alpha) = \alpha$$

α	z_α
.2	0.842
.15	1.036
.1	1.282
.075	1.440
.05	1.645
.04	1.751
.03	1.881
.025	1.960
.02	2.054
.01	2.326
.005	2.576
.001	3.090
.0005	3.291
.0001	3.719
.00005	3.891
.00001	4.265
.000005	4.417
.000001	4.753