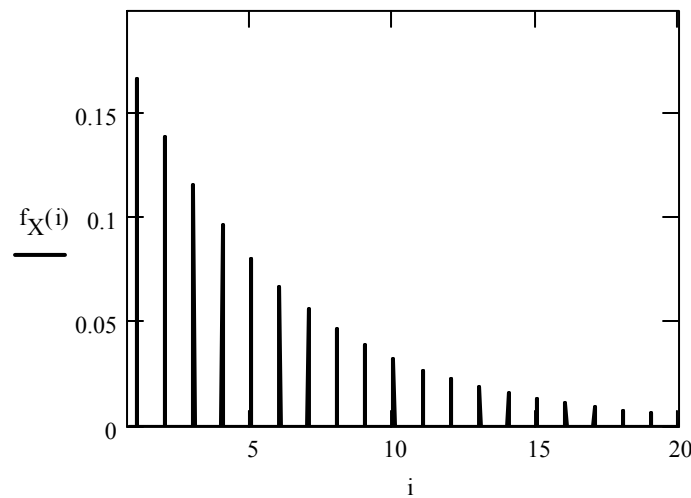


Probability and Random Variable (3)

- **The Geometric Random Variable**

- Suppose independent trials, each having a probability p of being a success, are performed.
- We define the geometric random variable (rv) X as the number of trials until the first success occurs.
- The pmf of X is defined as

$$f_X(i) = P\{X = i\} = (1 - p)^{i-1} p, \quad i = 1, 2, \dots$$



- Note that $f_X(i)$ defines a pmf since

$$\sum_{i=1}^{\infty} f_X(i) = p \sum_{i=1}^{\infty} (1 - p)^{i-1} = p \sum_{i=0}^{\infty} (1 - p)^i = p / [1 - (1 - p)] = 1.$$

- Let $q = 1 - p$. The expected value of X is

$$\begin{aligned}
 E[X] &= \sum_{i=1}^{\infty} i q^{i-1} p = p \sum_{i=1}^{\infty} \frac{d}{dq} (q^i) = p \frac{d}{dq} \left(\sum_{i=1}^{\infty} q^i \right) \\
 &= p \frac{d}{dq} \left(\frac{q}{1-q} \right) = \frac{p}{(1-q)^2} = \frac{1}{p}.
 \end{aligned}$$

➤ Similarly,

$$\begin{aligned}
 E[X^2] &= p \sum_{i=1}^{\infty} i^2 q^{i-1} = p \frac{(1+q)}{(1-q)^3} = \frac{1+q}{(1-q)^2} = \frac{2-p}{p^2}, \\
 \text{Var}[X] &= E[X^2] - (E[X])^2 = \frac{(2-p)}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.
 \end{aligned}$$

Example 12.

- When rolling a die repetitively, what is the probability that the first 6 appear on the sixth roll?
- Let X be the number of rolls until a 6 appears. Then, X is a geometric rv with parameter $p = 1/6$, and the desired probability is $P\{X=5\} = (5/6)^5(1/6) = 0.0667$.
- What is the expected number of rolls until a 6 appears?
- $E[X] = 1/p = 6$.

• Moment Generating Function (mgf)

➤ The mgf of a rv X is

$$\phi_X(t) = E[e^{tX}] = \begin{cases} \sum_{x_i} e^{tx_i} f_X(x_i) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx & \text{if } X \text{ is continuous} \end{cases}$$

Fact. The mgf has the following properties

(i) The mgf uniquely defines a rv .

$$(ii) E[X^n] = \left. \frac{d^n \phi_X(t)}{dt^n} \right|_{t=0};$$

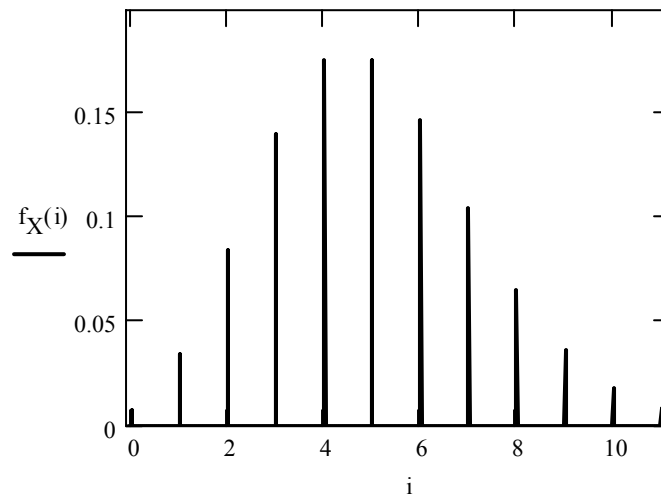
(iii) If X and Y are independent random variables, and $Z = X+Y$;

$$\text{then } \phi_Z(t) = \phi_{X+Y}(t) = \phi_X(t)\phi_Y(t).$$

- **The Poisson Random Variable**

➤ A rv, taking on values $0, 1, \dots$, is said to be a Poisson random variable with parameter $\lambda > 0$ if

$$f_X(i) = P\{X = i\} = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0, 1, \dots$$



➤ $f_X(i)$ defines a pmf since $\sum_{i=0}^{\infty} f_X(i) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = (e^{-\lambda})(e^{\lambda}) = 1.$

➤ The Poisson rv is a good model for demand, arrivals, and certain rare events.

- The mgf of X is

$$\phi_X(t) = \sum_{i=0}^{\infty} e^{ti} e^{-\lambda} \frac{\lambda^i}{i!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{(\lambda e^t)^i}{i!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)} .$$

- The first two moments and the variance of X are

$$E[X] = \left. \frac{d\phi_X(t)}{dt} \right|_{t=0} = \left. \lambda e^t e^{\lambda(e^t-1)} \right|_{t=0} = \lambda ,$$

$$E[X^2] = \left. \frac{d^2\phi_X(t)}{dt^2} \right|_{t=0} = \left. \lambda e^t (e^{\lambda(e^t-1)}) + \lambda e^t (\lambda e^t e^{\lambda(e^t-1)})^2 \right|_{t=0} = \lambda + \lambda^2 ,$$

$$Var[X] = E[X^2] - (E[X])^2 = \lambda .$$

- Let X_1 and X_2 be two independent Poisson rv's with means λ_1 and λ_2 . Then,

$$\phi_{X_1+X_2}(t) = \phi_{X_1}(t)\phi_{X_2}(t) = e^{\lambda_1(e^t-1)} \times e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)} ,$$

which implies that $X_1 + X_2$ is a Poisson rv with mean $\lambda_1 + \lambda_2$.

Example 13.

- The monthly demand for a certain airplane spare part of Fly High Airlines (FHA) fleet at Beirut airport is estimated to be a Poisson random variable with mean 0.5. Suppose that FHA will stock one spare part at the beginning of March. Once the part is used, a new part is ordered. The delivery lead time for a part is 2 months.
- What is the probability that the spare part will be used during March?
- Let X be the demand for the spare part. The desired probability is $P\{X \geq 1\} = 1 - e^{-\lambda} = 1 - e^{-0.5} = 0.393$.
- What is the probability that FHA will face a shortage on this part in March?

- The desired probability is $P\{X > 1\} = 1 - P\{X = 0\} - P\{X = 1\}$
 $= 1 - e^{-0.5} - 0.5e^{-0.5} = 0.09$.

Example 14.

- The number of people in a car entering Joe's Diner parking lot is a Poisson random variable with mean 2.5.
- If 10 cars are in the parking during lunch time, what is the probability that the number of customers having lunch is less than or equal to 20?
- The number of customers having lunch is a Poisson rv with mean $10 \times 2.5 = 25$. Then, the desired probability is

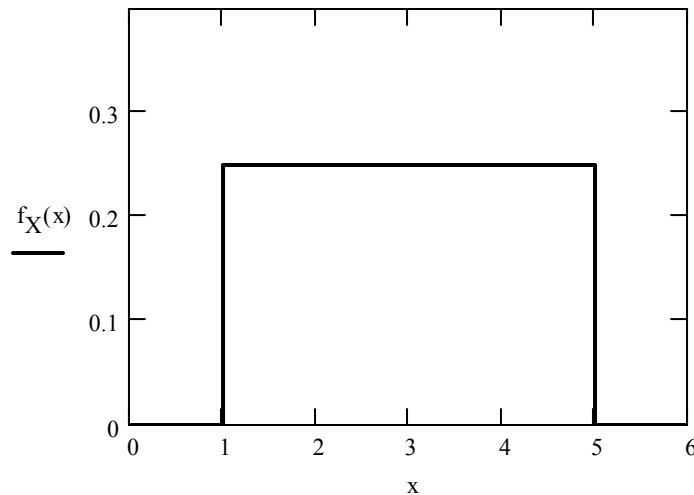
$$\sum_{i=0}^{20} e^{-25} \frac{25^i}{i!} = 0.185.$$

• The Uniform Random Variable

- A rv X that is equally like to be "near" any point of an interval (a, b) is said to have a uniform distribution.
- The pdf of X is

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a < x < b \\ 0, & \text{otherwise} \end{cases}$$

- Note that $f_X(x)$ defines a pdf since $\int_a^b f_X(x) dx = \int_a^b \frac{1}{b-a} dx = 1$.



➤ The cdf of X is

$$F_X(x) = \int_{-\infty}^x f_X(t) dt = \begin{cases} 0, & \text{if } x < a \\ \frac{x-a}{b-a}, & \text{if } a \leq x \leq b \\ 1, & \text{otherwise} \end{cases}$$

➤ The first two moments of X are

$$E[X] = \int_a^b x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2},$$

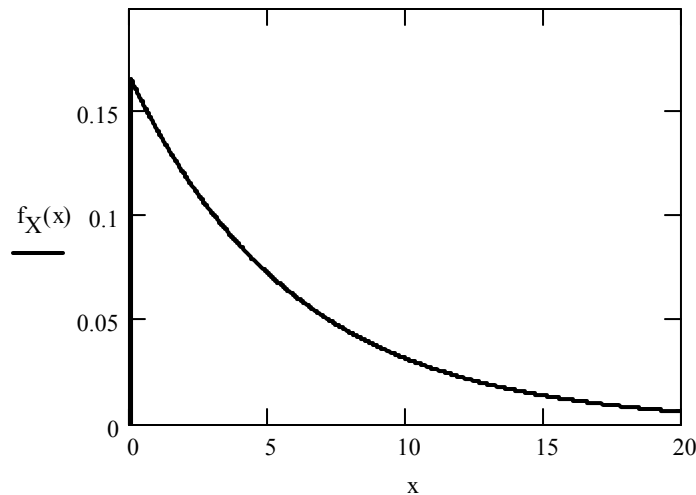
$$E[X^2] = \int_a^b x^2 f_X(x) dx = \int_a^b \frac{x^2}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}.$$

➤ The variance of X is $E[X^2] - (E[X])^2 = (b-a)^2 / 12$.

• The Exponential Random Variable

➤ An exponential rv with parameter λ is a rv whose pdf is

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$



- Note that $f_X(x)$ defines a pdf since

$$\int_0^{\infty} f_X(x) dx = \int_0^{\infty} \lambda e^{-\lambda x} dx = -e^{-\lambda x} \Big|_0^{\infty} = 1.$$

- The exponential rv is a good model for time between arrivals or time to failure of certain equipments.
- The cdf of X is

$$F_X(x) = \int_0^x f_X(t) dt = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x}, \quad x \geq 0.$$

- The mgf of X is

$$\phi_X(t) = \int_0^{\infty} e^{tx} f_X(x) dx = \int_0^{\infty} \lambda e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda-t} e^{-(\lambda-t)x} \Big|_0^{\infty} = \frac{\lambda}{\lambda-t}, \quad \text{for } t < \lambda.$$

- A useful property of the exponential distribution is that

$$\mathbf{P}\{X > x\} = e^{-\lambda x}.$$

➤ The first two moments and the variance of X are

$$E[X] = \left. \frac{d\phi_X(t)}{dt} \right|_{t=0} = \left. \frac{\lambda}{(\lambda-t)^2} \right|_{t=0} = \frac{1}{\lambda},$$

$$E[X^2] = \left. \frac{d^2\phi_X(t)}{dt^2} \right|_{t=0} = \left. \frac{2\lambda}{(\lambda-t)^3} \right|_{t=0} = \frac{2}{\lambda^2},$$

$$\text{Var}[X] = E[X^2] - (E[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Proposition. *The exponential distribution has the **memoryless property**. I.e., $P\{X > t+u \mid X > t\} = P\{X > u\}$.*

Proof.

$$P\{X > t+u \mid X > t\} = \frac{P\{X > t+u, X > t\}}{P\{X > t\}} = \frac{P\{X > t+u\}}{P\{X > t\}}$$

$$= \frac{e^{-\lambda(t+u)}}{e^{-\lambda t}} = e^{-\lambda u} = P\{X > u\}.$$

➤ The memoryless property allows developing tractable analytical models with the exponential distribution. It makes the exponential distribution very popular in modeling.

Fact. *Let X_1 and X_2 be two independent exponential random variables with parameters λ_1 and λ_2 . Let $X = \min(X_1, X_2)$. Then, X is an exponential random variable with parameter $\lambda_1 + \lambda_2$.*

Proof. $P\{X > x\} = P\{X_1 > x, X_2 > x\} = P\{X_1 > x\}P\{X_2 > x\}$
 $= e^{-\lambda_1 x} e^{-\lambda_2 x} = e^{-(\lambda_1 + \lambda_2)x}$.

Example 15.

- The amount of time one spends in the bank is exponentially distributed with mean 10 minutes.
- A customer arrives at 1:00 PM. What is the probability that the customer will be in the bank at 1:15 PM?
- Let X be the time the customers spends in the bank. Then, X is exponentially distributed with parameter $\lambda = 1/10$. The desired probability is $P\{X > 15\} = e^{-15\lambda} = e^{-15/10} = 0.223$.
- It is now 1:20 PM and the customer is still in the bank? What is the probability that the customer will be in the bank at 1:35 PM?
- 0.223 (by the memoryless property).

Example 16.

- Two light bulbs have iid lifetimes which are exponentially distributed with mean 1000 hours. Bulb 1 is just installed, while Bulb 2 has been operational for 500 hours.
- What is the probability that Bulb 1 fails after 500 hours?
 $1 - e^{-500/1000} = 0.393$.
- What is the probability that Bulb 2 fails after 500 hours?
0.393 (by the memoryless property).

Example 17.

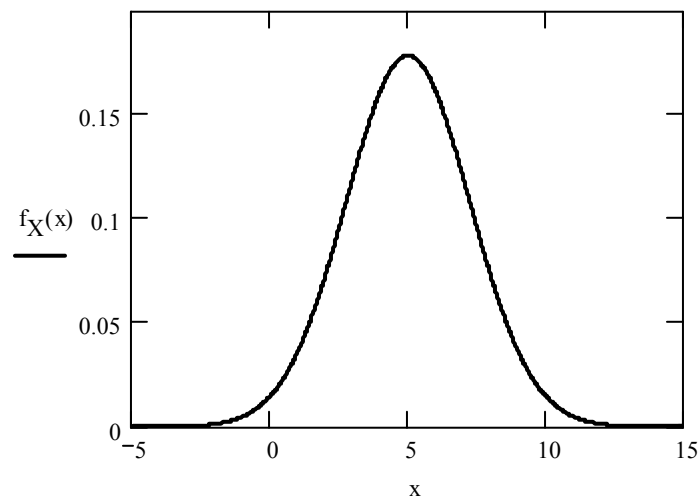
- A post office is run by two clerks. Mr. Smith enters the office and finds the two clerks busy serving Mr. Jones and Mr. Brown. The amount of time a clerk spends with a customer is exponentially distributed with mean 20 minutes.

- What is the probability that, out of the three customers Mr. Smith is the last to leave the post office?
- 1/2. (When one of the clerks becomes free, Mr. Smith starts service. Then, by the memoryless property Mr. Smith is equally likely to finish first as the customer served by the other clerk.)
- What is the probability that Mr. Smith waits for service more than 30 minutes?
- Mr. Smith will start service when either Mr. Jones or Mr. Brown finishes service. Then, the waiting time of Mr. Smith, X , is exponential with parameter $1/20 + 1/20 = 1/10$. The desired probability is therefore $P\{X > 30\} = e^{-30/10} = 0.05$.

- **The Normal Random Variable**

- We say that a random variable X is a normal rv with parameters μ and $\sigma > 0$ if it has the following pdf:

$$f_X(x) = \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma}, x \in (-\infty, \infty).$$



➤ Note that $f_X(x)$ defines a pdf. With a change of variable

$z = (x - \mu)/\sigma$ and using the fact that $\int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}$,

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_{-\infty}^{\infty} \frac{e^{-(x-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz = 1.$$

➤ The normal rv is a good model for quantities that can be seen as sums or averages of a large number of rv's.

➤ The cdf of X , $F_X(x) = \int_{-\infty}^x f_X(t) dt$, has no closed-form.

Fact. If X is a normal rv, then $Z = (X - \mu)/\sigma$ is a “standard normal r.v.” with parameters 0 and 1.

Proof. Note that

$$P\{Z < z\} = P\left\{\frac{X - \mu}{\sigma} < z\right\} = P\{X < \mu + \sigma z\} = \int_{-\infty}^{\mu + \sigma z} \frac{e^{-(t-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma} dt.$$

Let $u = (t - \mu)/\sigma$, then $P\{Z < z\} = \int_{-\infty}^z \frac{e^{-u^2/2}}{\sqrt{2\pi}} du$, which is the cdf of

the standard normal.

➤ This fact implies that $X = \mu + \sigma Z$.

➤ The mgf of Z (the standard normal r.v.) is

$$\begin{aligned} \phi_Z(t) &= \int_{-\infty}^{\infty} e^{tz} f_Z(z) dz = \int_{-\infty}^{\infty} \frac{e^{(2tz - z^2)/2}}{\sqrt{2\pi}} dz \\ &= \frac{e^{t^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(t-z)^2/2} dz = e^{t^2/2}. \end{aligned}$$

➤ The mgf of X is then

$$\phi_X(t) = E[e^{tX}] = E[e^{t(\mu + \sigma Z)}] = e^{t\mu} E[e^{t\sigma Z}] = e^{t\mu} e^{t^2\sigma^2/2} = e^{\sigma^2 t^2/2 + t\mu}.$$

➤ The first two moments of X are

$$E[X] = \left. \frac{d\phi_X(t)}{dt} \right|_{t=0} = (\mu + \sigma^2 t) e^{(\mu + \sigma^2 t^2/2)} \Big|_{t=0} = \mu,$$

$$E[X^2] = \left. \frac{d^2\phi_X(t)}{dt^2} \right|_{t=0} = \sigma^2 e^{(\mu + \sigma^2 t^2/2)} + (\mu + \sigma^2 t)^2 e^{(\mu + \sigma^2 t^2/2)} \Big|_{t=0} = \sigma^2 + \mu^2,$$

$$\text{Var}[X] = E[X^2] - (E[X])^2 = \sigma^2.$$

➤ The cdf of X , $F_X(x)$, is evaluated through the cdf of Z , $F_Z(z)$, which is often tabulated (see next page),

$$P\{X < x\} = P\left\{Z < \frac{x - \mu}{\sigma}\right\} \Rightarrow F_X(x) = F_Z\left(\frac{x - \mu}{\sigma}\right).$$

Theorem (central limit theorem). *If X_i , $i = 1, 2, \dots, n$, are iid rv's with mean μ and variance σ^2 . Then, for n large*

enough, $\sum_{i=1}^n X_i$ is normally distributed with mean $n\mu$ and

variance $n\sigma$.

Example 18.

➤ The height of an AUB male student is a normal rv with mean 170 cm and standard deviation 8 cm.

➤ What is the probability that the height of an AUB student is less than 180 cm?

➤ Let X be the height of the student. Then, the desired probability is

$$P\{X < 180\} = P\{Z < (180 - 170)/8\} = P\{Z < 1.25\} = 0.894.$$

Example 19.

- What is the probability that a normal random variable, $X(\mu, \sigma)$, takes on values in the interval $(\mu - 2\sigma, \mu + 2\sigma)$?

$$\begin{aligned} P\{\mu - 2\sigma < X < \mu + 2\sigma\} &= P\{-2 < Z < 2\} = P\{Z < 2\} - P\{Z < -2\} \\ &= P\{Z < 2\} - (1 - P\{Z < 2\}) \\ &= 2P\{Z < 2\} - 1 = 2 \times 0.9772 - 1 = 0.954. \end{aligned}$$

- What is the probability that a normal random variable, $X(\mu, \sigma)$, takes on values in the interval $(\mu - 3\sigma, \mu + 3\sigma)$?
- Similar, $= 2P\{Z < 3\} - 1 = 2 \times 0.9987 - 1 = 0.997$.
- Insight?
- For a normal distributed rv, most values fall within two standard deviations of the mean. Almost *all* values fall within three standard deviations of the mean (partial motivating for the 6-Sigma jargon.)
- Practical implications.
- Standard deviation \approx range /4.
- With Mean \approx Median, there is no need for a calculator to do two practical two-moments stats, e.g. when grading exams.
- E.g. ENMG 602, Fall 2008 final exam grades. The number of students in that class was 42.

Mean	85
Standard deviation	8
Median	87
Max	99
Min	71
Range	28

Area $P\{Z \leq z\}$ under the standard normal curve to the left of z

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Examples. $P\{Z \leq 0.41\} = 0.6591$,
 $P\{Z < -0.54\} = 1 - P\{Z \leq 0.54\} = 1 - 0.7054 = 0.2946$,
 $P\{Z > 0\} = 1 - P\{Z \leq 0\} = 1 - 0.5000 = 0.5000$.