

Examiner: Xiangfeng Yang (Tel: 013 28 57 88).

Things allowed (Hjälpmedel): a calculator.

Scores rating (Betygsgränser): 8-11 points giving rate 3; 11.5-14.5 points giving rate 4; 15-18 points giving rate 5.

Write down all necessary steps in solutions in order to receive as many points as possible.

1 (3 points)

Suppose that a box contains 6 red balls and 8 blue balls. One now randomly takes 2 balls at once from the box. Define the following three events:

$$A = \{2 \text{ balls have same color}\},$$

$$B = \{1 \text{ red ball and 1 blue ball}\},$$

$$C = \{\text{at least 1 red ball}\}.$$

(1.1) (1p) Find the probability $P(A)$.

(1.2) (1p) Find the probability $P(A \cap C)$.

(1.3) (1p) Find the conditional probability $P(B | C)$.

Solution. (1.1)

$$P(A) = P(2 \text{ balls are red}) + P(2 \text{ balls are blue}) = \frac{\binom{6}{2}}{\binom{14}{2}} + \frac{\binom{8}{2}}{\binom{14}{2}} = \frac{15}{91} + \frac{28}{91} = \frac{43}{91} = 0.4725.$$

(1.2)

$$P(A \cap C) = P(2 \text{ balls are red}) = \frac{\binom{6}{2}}{\binom{14}{2}} = \frac{15}{91} = 0.1648.$$

(1.3)

$$\begin{aligned} P(B | C) &= \frac{P(B \cap C)}{P(C)} = \frac{P(1 \text{ red ball and 1 blue ball})}{1 - P(\text{no red ball})} = \frac{\frac{\binom{6}{1} \cdot \binom{8}{1}}{\binom{14}{2}}}{1 - \frac{\binom{8}{2}}{\binom{14}{2}}} \\ &= \frac{\frac{48}{91}}{1 - \frac{28}{91}} = \frac{48}{63} = \frac{16}{21} = 0.7619. \end{aligned}$$

□

2 (3 points)

During the period 2024.1.1-2024.5.27, suppose that 25% of all LiU students have visited IKEA Linköping.

(2.1) (1p) Among 10 randomly selected LiU students, what is the probability that there are at most 5 who have visited IKEA Linköping during the period 2024.1.1-2024.5.27?

(2.2) (2p) Among 100 randomly selected LiU students, what is the probability that there are at most 35 who have visited IKEA Linköping during the period 2024.1.1-2024.5.27?

Solution. (2.1) Let X be the number of LiU students who have visited IKEA Linköping during the period 2024.1.1-2024.5.27 among these 10 randomly selected persons, then $X \sim \text{Bin}(10, 25\%)$. Therefore

$$P(\text{at most } 5) = P(X \leq 5) = (\text{Binomial table}) = 0.9803.$$

(2.2) Let Y be the number of LiU students who have visited IKEA Linköping during the period 2024.1.1-2024.5.27 among these 100 randomly selected persons, then $Y \sim \text{Bin}(100, 25\%)$. Normal approximation to Binomial shall be used

$$\begin{aligned} P(\text{at most } 35) &= P(Y \leq 35) = P(N(100 \cdot 25\%, 100 \cdot 25\% \cdot (1 - 25\%)) \leq 35) \\ &= P(N(25, 18.75) \leq 35) = P(N(0, 1) \leq \frac{35 - 25}{\sqrt{18.75}}) \\ &= P(N(0, 1) \leq 2.31) = \Phi(2.31) = 0.9896. \end{aligned}$$

□

3 (3 points)

Lars is table tennis player, and he spends some time everyday for training. Let X denote the time (in hours) Lars spends in one day for training. Assume that X is a continuous random variable with a probability density function:

$$f(x) = \begin{cases} \frac{x}{8}, & \text{if } 0 \leq x \leq 4, \\ 0, & \text{otherwise.} \end{cases}$$

(3.1) (1p) Find the mean $\mu = E(X)$ and variance $\sigma^2 = V(X)$ of X .

(3.2) (2p) Let X_1, X_2, \dots, X_{100} denote the times Lars spends for training in 100 days, and assume that X_1, X_2, \dots, X_{100} are independent. Find the probability that the total training time in these 100 days is less than 280 hours, i.e. $P(X_1 + X_2 + \dots + X_{100} < 280)$.

Solution. (3.1)

$$\begin{aligned} \mu = E(X) &= \int_0^4 x f(x) dx = \frac{1}{8} \int_0^4 x^2 dx = \frac{1}{8 \cdot 3} [x^3]_0^4 = \frac{8}{3} = 2.6667, \\ \sigma^2 = V(X) &= E(X^2) - \mu^2 = \int_0^4 x^2 f(x) dx - \left(\frac{8}{3}\right)^2 = \frac{1}{8 \cdot 4} [x^4]_0^4 - \left(\frac{8}{3}\right)^2 = 8 - \frac{64}{9} = \frac{8}{9} = 0.8889. \end{aligned}$$

(3.2)

$$\begin{aligned} P(X_1 + X_2 + \dots + X_{100} < 280) &= P(\bar{X} < 2.8) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{2.8 - \mu}{\sigma/\sqrt{n}}\right) \\ &= P\left(N(0, 1) < \frac{2.8 - \frac{8}{3}}{\sqrt{\frac{8}{9}/\sqrt{100}}}\right) = P\left(N(0, 1) < \frac{28 - \frac{80}{3}}{\sqrt{\frac{8}{9}}}\right) = P\left(N(0, 1) < \frac{28 \cdot 3 - 80}{\sqrt{8}}\right) \\ &= P(N(0, 1) < 1.41) = \Phi(1.41) = 0.9207. \end{aligned}$$

□

4 (3 points)

(4.1) (1p) A discrete population X has the following distribution

X	1	5	8
$p(x)$	$(1 - \theta)/3$	$1/3$	$(1 + \theta)/3$

where θ is unknown. To estimate θ , a sample $\{x_1, x_2, \dots, x_n\}$ is taken from this population. Use the method of moments to find a point estimate $\hat{\theta}_{MM}$ of θ .

(4.2) (2p) Another population Y is a continuous random variable with a probability density function

$$f(y) = \frac{a}{y^{a+1}}, \quad \text{for } y \geq 1,$$

where $a > 1$ is unknown. A sample $\{y_1, y_2, \dots, y_n\}$ is taken from this population. Use the maximum-likelihood method to find a point estimate \hat{a}_{ML} of a .

Solution. (4.1) The mean $E(X)$ of the population can be computed as

$$E(X) = 1 \cdot \frac{1-\theta}{3} + 5 \cdot \frac{1}{3} + 8 \cdot \frac{1+\theta}{3} = \frac{1}{3}(14 + 7\theta).$$

It is then from the equation $E(X) = \bar{x}$ that

$$\frac{1}{3}(14 + 7\theta) = \bar{x} \Rightarrow \hat{\theta}_{MM} = \frac{3\bar{x} - 14}{7}.$$

(4.2) The likelihood function is

$$L(a) = f(y_1) \cdot f(y_2) \cdot \dots \cdot f(y_n) = \frac{a}{y_1^{a+1}} \cdot \frac{a}{y_2^{a+1}} \cdot \dots \cdot \frac{a}{y_n^{a+1}} = \frac{a^n}{(y_1 y_2 \dots y_n)^{a+1}}.$$

The log likelihood function is then

$$\ln L(a) = n \ln(a) - (a+1) \ln(y_1 y_2 \dots y_n).$$

The equation $0 = \ln' L(a)$ gives that

$$0 = \frac{n}{a} - \ln(y_1 y_2 \dots y_n),$$

which implies that

$$\hat{a}_{ML} = \frac{n}{\ln(y_1 y_2 \dots y_n)}.$$

□

5 (3 points)

One wants to study waiting times for meals in 2 fast food restaurants: McDonald's Tornby and Burger King Tornby.

(5.1) (1p) For McDonald's Tornby, suppose that the distribution X of waiting times is $X \sim N(\mu_x, \sigma^2)$ where μ_x is the average waiting time. A sample is taken from X with $n = 16$, $\bar{x} = 4.2$ and $s_x = 2.2$. Construct a 95% (two-sided) confidence interval I_{μ_x} of μ_x .

(5.2) (2p) For Burger King Tornby, suppose that the distribution Y of waiting times is $Y \sim Exp(\mu_y)$ where μ_y (the average waiting time) is the mean of Y . A sample is taken from Y with $m = 36$ and $\bar{y} = 5.5$. Construct a 95% (two-sided) confidence interval I_{μ_y} of μ_y .

Solution. (5.1) This is Case 1.3, and the confidence interval is

$$I_{\mu_x} = \bar{x} \mp t_{\alpha/2}(n-1) \frac{s_x}{\sqrt{n}} = 4.2 \mp t_{0.025}(15) \frac{2.2}{\sqrt{16}} = 4.2 \mp 2.13 \frac{2.2}{\sqrt{16}} = 4.2 \mp 1.1715 = (3.0285, 5.3715).$$

(5.2). This is Case 1.2, and the confidence interval is (because σ is unknown)

$$I_{\mu_y} = \bar{y} \mp z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{m}} = 5.5 \mp 1.96 \frac{5.5}{\sqrt{36}} = 5.5 \mp 1.7967 = (3.7033, 7.2967).$$

where we have used the fact that for exponential random variable $\sigma = \mu$ implying that $\hat{\sigma} = \hat{\mu} = \bar{x}$.

□

6 (3 points)

Suppose that a population $X \sim N(\mu, 3^2)$, and a sample is taken from this population with $n = 9$, $\bar{x} = 6.6$ and $s_x = 3.3$. We want to test the following hypotheses with a significance level $\alpha = 5\%$,

$$H_0 : \mu = 6 \quad \text{against} \quad H_a : \mu > 6,$$

(6.1) (1p) Is H_0 rejected based on TS (observed test statistic under H_0) and C (rejection region)?

(6.2) (1p) Is H_0 rejected based on p -value?

(6.3) (1p) Find the power when the true value is $\mu = 7$.

Solution. (6.1) Since here σ^2 is known, one has HT-1 (case 1.1) and

$$TS = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{6.6 - 6}{3/\sqrt{9}} = 0.6, \quad C = (z_\alpha, \infty) = (1.64, \infty),$$

and $TS \notin C$, which implies that H_0 is not rejected.

(6.2) In this case, one can calculate the p -value as

$$p\text{-value} = P(UL \geq TS) = P(N(0,1) \geq 0.6) = 1 - P(N(0,1) < 0.6) = 1 - \Phi(0.6) = 1 - 0.7257 = 0.2743.$$

Since $p\text{-value} \not\leq \alpha$, we do not reject H_0 .

(6.3) The power is computed as follows

$$\begin{aligned} h(7) &= P(H_0 \text{ is rejected when } H_0 \text{ is false and } \mu = 7) = P(TS \in C, \text{ when } \mu = 7) \\ &= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > 1.64, \text{ when } \mu = 7\right) \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > 1.64 + \frac{(\mu_0 - \mu)}{\sigma/\sqrt{n}}, \text{ when } \mu = 7\right) \\ &= P(N(0,1) > 0.64) = 1 - P(N(0,1) \leq 0.64) = 1 - \Phi(0.64) = 1 - 0.7389 = 0.2611. \end{aligned}$$

□

1. Basic probability

(1.1) Conditional probability $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

(1.2) Total probability $P(B) = \sum_{i=1}^k P(B|A_i)P(A_i)$ where $\{A_i\}$ are disjoint and $\cup_{i=1}^k A_i = S$.

(1.3) Bayes' Theorem $P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^k P(B|A_i)P(A_i)}$ where $\{A_i\}$ are in (1.2).

2. Random variables (r.v.s)

(2.1) Discrete r.v. X has a pmf $p(x) = P(X = x)$ satisfying $p(x) \geq 0$ and $\sum p(x_i) = 1$,

$$\begin{array}{c|cccc} X & x_1 & x_2 & \cdots & x_n & \cdots \\ \hline p(x) & p(x_1) & p(x_2) & \cdots & p(x_n) & \cdots \end{array}$$

Expectation (or *Expected value* or *mean*) $\mu_X = E(X) = \sum x_i p(x_i)$;

Variance $\sigma_X^2 = V(X) = E(X - \mu_X)^2 = E(X^2) - \mu_X^2 = \sum x_i^2 p(x_i) - (\sum x_i p(x_i))^2$.

(2.2) Continuous r.v. X has a pdf $f(x)$ satisfying $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x) dx = 1$,

$$P(a < X < b) = \int_a^b f(x) dx.$$

Expectation (or *Expected value* or *mean*) $\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) dx$;

Variance $\sigma_X^2 = V(X) = E(X - \mu_X)^2 = E(X^2) - \mu_X^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - (\int_{-\infty}^{\infty} x f(x) dx)^2$.

(2.3) Cumulative distribution function (cdf) of a r.v. X is $F(x) = P(X \leq x)$.

(2.4) X and Y are r.v.s, a, b and c are scalars, then

$$E(aX + bY + c) = aE(X) + bE(Y) + c,$$

$$V(aX + bY + c) = a^2V(X) + b^2V(Y) + 2ab \operatorname{cov}(X, Y),$$

$$E(g(X, Y)) = \begin{cases} \sum_{i,j} g(x_i, y_j) \cdot p(x_i, y_j), & \text{for discrete } (X, Y), \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot f(x, y) dx dy, & \text{for continuous } (X, Y). \end{cases}$$

(2.5) • Discrete r.v. (X, Y) has a joint pmf $p(x, y)$ satisfying $p(x, y) \geq 0$ and $\sum_{x_i} \sum_{y_j} p(x_i, y_j) = 1$.

The *marginal pmf* of X is $p_X(x) = \sum_y p(x, y)$;

The *marginal pmf* of Y is $p_Y(y) = \sum_x p(x, y)$;

X and Y are *independent* if $p(x, y) = p_X(x) \cdot p_Y(y)$.

• Continuous r.v. (X, Y) has a joint pdf $p(x, y)$ satisfying $f(x, y) \geq 0$ and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

The *marginal pdf* of X is $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$;

The *marginal pdf* of Y is $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$;

X and Y are *independent* if $f(x, y) = f_X(x) \cdot f_Y(y)$.

3. Several special r.v.s

(3.1) $X \sim \operatorname{Bin}(n, p)$ has a pmf $p(x) = P(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$, $x = 0, 1, 2, \dots, n$.

$$E(X) = n \cdot p, \quad V(X) = n \cdot p \cdot (1-p).$$

(3.2) $X \sim \operatorname{Po}(\lambda)$ has a pmf $p(x) = P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$, $x = 0, 1, 2, \dots$
 $E(X) = \lambda, \quad V(X) = \lambda$.

(3.3) $X \sim \operatorname{Hypergeometric}$ has a pmf $p(x) = P(X = x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$.

(3.4) $X \sim \operatorname{Exp}(\lambda)$ has a pdf

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

(3.5) $E(X) = \frac{1}{\lambda}, \quad V(X) = \left(\frac{1}{\lambda}\right)^2$.

$X \sim N(\mu, \sigma^2)$ has a pdf

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

(3.6) $E(X) = \mu, \quad V(X) = \sigma^2$.

$X \sim U(a, b)$ has a pdf

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

$$E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}.$$

4. Central Limit Theorem (CLT)

Suppose that a population has mean $= \mu$ and variance $= \sigma^2$. A random sample $\{X_1, X_2, \dots, X_n\}$ from this population is given. Then for large $n \geq 30$,

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1). \tag{1}$$

• If the population is normal, then (1) holds for any n .

• Note that $\mu = E(\bar{X})$ and $(\sigma/\sqrt{n})^2 = V(\bar{X})$.

5. Several notations in statistics

(5.1) Sample mean: $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \sum \frac{X_i}{n}; \quad \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \sum \frac{x_i}{n}$.

(5.2) Sample variance:

$$S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1} = \frac{1}{n-1} \left(\sum X_i^2 - \frac{(\sum X_i)^2}{n} \right); \quad s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{1}{n-1} \left(\sum x_i^2 - \frac{(\sum x_i)^2}{n} \right).$$

• Capital letters \bar{X} and S^2 refer to the objects based on random sample (therefore they are in general r.v.s), while small letters \bar{x} and s^2 are the objects based on observations (so they are scalars).

(5.3) A point estimator of θ obtained by Methods of Moments is denoted as $\hat{\theta}_{MM}$.

(5.4) A point estimator of θ obtained by Maximum Likelihood method is denoted as $\hat{\theta}_{ML}$.

6. Confidence Interval (CI)

In this course, three types of confidence intervals are studied depending on the unknown population parameter(s): CI-1 (confidence intervals for population mean(s)), CI-2 (confidence intervals for population variance(s)), and CI-3 (confidence intervals for population proportion(s)).

CI-1: (1 - α) CI of a population mean μ

case 1.1 (any n) If population $X \sim N(\mu, \sigma^2)$ and σ^2 is known, then $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$ and

$$I_\mu = (\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) := \bar{x} \mp z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}.$$

case 1.2 (n ≥ 30) For any population X , it holds that $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0, 1)$ and

$$I_\mu = \bar{x} \mp z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \text{ or } I_\mu = \bar{x} \mp z_{\alpha/2} \cdot \frac{\hat{\sigma}}{\sqrt{n}}.$$

case 1.3 (any n) If population $X \sim N(\mu, \sigma^2)$ and σ^2 is unknown, then $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim T(n-1)$ and

$$I_\mu = \bar{x} \mp t_{\alpha/2}(n-1) \cdot \frac{s}{\sqrt{n}}.$$

CI-1': (1 - α) CI of the difference of two population means $\mu_X - \mu_Y$

case 1.1' (any n_1, n_2) If independent populations $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$, and σ_X^2, σ_Y^2 are known, then

$$\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}} \sim N(0, 1), \text{ and } I_{\mu_X-\mu_Y} = (\bar{x}-\bar{y}) \mp z_{\alpha/2} \cdot \sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}.$$

case 1.2' ($n_1, n_2 \geq 30$) For any independent populations X and Y , it holds that

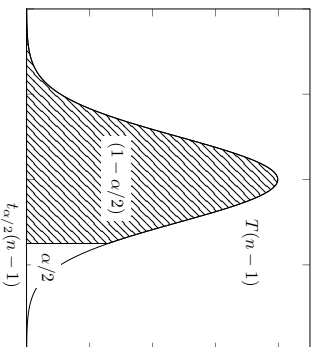
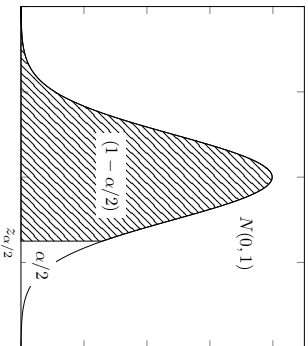
$$\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}} \sim N(0, 1) \text{ and}$$

$$I_{\mu_X-\mu_Y} = (\bar{x}-\bar{y}) \mp z_{\alpha/2} \cdot \sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}} \text{ or } I_{\mu_X-\mu_Y} = (\bar{x}-\bar{y}) \mp z_{\alpha/2} \cdot \sqrt{\frac{\hat{\sigma}_X^2}{n_1} + \frac{\hat{\sigma}_Y^2}{n_2}}.$$

case 1.3' (any n_1, n_2) If independent populations $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$, where σ_X^2, σ_Y^2 are unknown but $\sigma_X^2 = \sigma_Y^2$, then

$$\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim T(n_1+n_2-2), \text{ where } S^2 = \frac{(n_1-1)S_X^2 + (n_2-1)S_Y^2}{n_1+n_2-2}, \text{ and}$$

$$I_{\mu_X-\mu_Y} = (\bar{x}-\bar{y}) \mp t_{\alpha/2}(n_1+n_2-2) \cdot s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$



CI-2: (1 - α) CI of population variance(s) σ^2

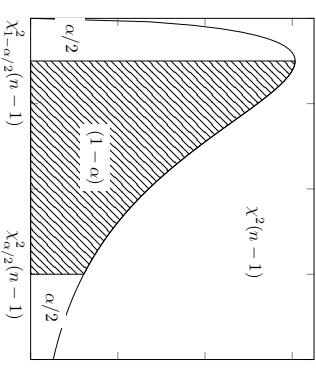
• If a population $X \sim N(\mu, \sigma^2)$ and σ^2 is unknown, then $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, and

$$I_{\sigma^2} = \left(\frac{(n-1)s^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)s^2}{\chi_{1-\alpha/2}^2(n-1)} \right).$$

• If two independent populations $X \sim N(\mu_X, \sigma^2)$ and $Y \sim N(\mu_Y, \sigma^2)$, and σ^2 is unknown, then $\frac{(n_1+n_2-2)S^2}{\sigma^2} \sim \chi^2(n_1+n_2-2)$, and

$$I_{\sigma^2} = \left(\frac{(n_1+n_2-2)s^2}{\chi_{\alpha/2}^2(n_1+n_2-2)}, \frac{(n_1+n_2-2)s^2}{\chi_{1-\alpha/2}^2(n_1+n_2-2)} \right),$$

where $S^2 = \frac{(n_1-1)S_X^2 + (n_2-1)S_Y^2}{n_1+n_2-2}$.



CI-3: (1 - α) CI of population proportion(s)

• If a (large) population has an unknown proportion p , then $\frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \sim N(0, 1)$ if $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$ with $\hat{p} = x/n$, and $I_p = \hat{p} \mp z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

• If two independent (large) populations have unknown proportions p_1 and p_2 , then

$$\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}} \sim N(0, 1)$$

if $n_i\hat{p}_i \geq 10$ and $n_i(1-\hat{p}_i) \geq 10$ for $i = 1, 2$, and $I_{p_1-p_2} = (\hat{p}_1-\hat{p}_2) \mp z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$.

7. Hypothesis Test (HT)

	H_0 is true	H_0 is false and $\theta = \theta_1$
reject H_0	(type I error or significance level) α	(power) $h(\theta_1)$
don't reject H_0	$1 - \alpha$	(type II error) $\beta(\theta_1) = 1 - h(\theta_1)$

reject $H_0 \Leftrightarrow TS \in C \Leftrightarrow p\text{-value} < \alpha$

χ^2 tests for populations (non-parametric tests)

Suppose that for a random sample of a population X , the n elements of it are classified into k disjoint groups $A_i, 1 \leq i \leq k$. For each group $A_i, 1 \leq i \leq k$, suppose that there are $N_i, 1 \leq i \leq k$ elements inside. Let $p_i = P(A_i)$ assuming a given distribution of X . Note that $p_1 + p_2 + \dots + p_k = 1$ and $N_1 + N_2 + \dots + N_k = n$. One wants to test the hypotheses

$$H_0 : P(A_i) = p_i, \quad 1 \leq i \leq k, \quad H_a : P(A_i) \neq p_i \text{ for some } 1 \leq i \leq k.$$

If n is large in the sense that $np_i \geq 5$ for all $1 \leq i \leq k$, then the test statistic is

$$\sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i} \approx \chi^2(k-1).$$

Therefore the observation of the test statistic is

$$TS = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}, \text{ where } n_i \text{ is the observation of } N_i, 1 \leq i \leq k.$$

For the critical region C , one can take (note that if H_0 is true, then TS should be close to zero)

$$C = (\chi^2_{\alpha}(k-1), \infty).$$

The conclusion would be $TS \in C \iff H_0$ is rejected.

8. Linear and logistic regression

(Multiple) linear regression: $Y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon$, $\varepsilon \sim N(0, \sigma^2)$.

- Y : response variable (which is normal r.v.), $\{x_1, \dots, x_k\}$: predictors (which are scalars).
- sample: $\{(x_{11}, \dots, x_{1k}; y_1), (x_{21}, \dots, x_{2k}; y_2), \dots, (x_{n1}, \dots, x_{nk}; y_n)\}$.
- how to estimate $\beta_j \approx \hat{\beta}_j$: least square method, that is, to minimize $\sum_{i=1}^n (y_i - \hat{y}_i)^2$, where the estimated (multiple) linear regression line \hat{y} is

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k.$$

- $\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim T(n-k-1)$, this helps determine whether or not the real $\beta_j = 0$?
- $\sigma^2 \approx \frac{SSE}{n-k-1}$, this gives an estimation of the size of the error.
- $R^2 = \frac{SSR}{SSY}$, this gives how well the model is (if $R^2 \approx 1$, then the model fits the sample very well).
- How to test $\beta_1 = \dots = \beta_k = 0$? Use the random variable $\frac{SSR/k}{SSE/(n-k-1)} \sim F(k, n-k-1)$.

Logistic regression: Let Y can only take 0 or 1 with $P(Y=1) = p$ and $P(Y=0) = 1-p$.

$$E(Y) = p(x_1, \dots, x_k) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}.$$

- Y : response variable (which is Bernoulli r.v. $P(Y=1) = p$ and $P(Y=0) = 1-p$, so $E(Y) = p$), $\{x_1, \dots, x_k\}$: predictors (which are scalars).
- sample: $\{(x_{11}, \dots, x_{1k}; y_1), (x_{21}, \dots, x_{2k}; y_2), \dots, (x_{n1}, \dots, x_{nk}; y_n)\}$.
- how to estimate $\beta_j \approx \hat{\beta}_j$: maximal likelihood method (maximize $\prod_{i=1}^n p(x_{i1}, \dots, x_{ik})^{y_i} (1 - p(x_{i1}, \dots, x_{ik}))^{1-y_i}$).
- $\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \approx N(0, 1)$ for large $n \geq 30$, this helps determine whether or not the real $\beta_j = 0$?
- Classification of a new object $Y(x_1, \dots, x_k)$ as 1 or 0 according

$$Y(x_1, \dots, x_k) = \begin{cases} 1, & \text{if } \hat{p}(x_1, \dots, x_k) \geq 0.5, \\ 0, & \text{if } \hat{p}(x_1, \dots, x_k) < 0.5, \end{cases}$$

where the estimated logit function $\hat{p}(x_1, \dots, x_k)$ is

$$\hat{p}(x_1, \dots, x_k) = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k}}.$$

9. Tables

(9.1) Table for $N(0, 1)$ standard normal random variable $\Phi(x) = P(N(0, 1) \leq x)$, $x \geq 0$.

There is an important relation $\Phi(-x) = 1 - \Phi(x)$, $x \geq 0$.

x	0	1	2	3	4	5	6	7	8	9
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	0.9564	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	0.9993	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	0.9997	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998	0.9998	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	0.9998	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
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

(9.2) Table for $T(f)$ random variable $F(x) = P(T(f) \leq x)$,

where f is a parameter called 'degrees of freedom'.

f	0.75	0.90	0.95	0.975	0.99	0.995	0.9975	0.9995
1	1.00	3.08	6.31	12.71	31.82	63.66	127.32	636.62
2	0.82	1.89	2.92	4.30	6.96	9.92	14.09	31.60
3	0.76	1.64	2.35	3.18	4.54	5.84	7.45	12.92
4	0.74	1.53	2.13	2.78	3.75	4.60	5.60	8.61
5	0.73	1.48	2.02	2.57	3.36	4.03	4.77	6.87
6	0.72	1.44	1.94	2.45	3.14	3.71	4.32	5.96
7	0.71	1.41	1.89	2.36	3.00	3.50	4.03	5.41
8	0.71	1.40	1.86	2.31	2.90	3.36	3.83	5.04
9	0.70	1.38	1.83	2.26	2.82	3.25	3.69	4.78
10	0.70	1.37	1.81	2.23	2.76	3.17	3.58	4.59
11	0.70	1.36	1.80	2.20	2.72	3.11	3.50	4.44
12	0.70	1.36	1.78	2.18	2.68	3.05	3.43	4.32
13	0.69	1.35	1.77	2.16	2.65	3.01	3.37	4.22
14	0.69	1.35	1.76	2.14	2.62	2.98	3.33	4.14
15	0.69	1.34	1.75	2.13	2.60	2.95	3.29	4.07
16	0.69	1.34	1.75	2.12	2.58	2.92	3.25	4.01
17	0.69	1.33	1.74	2.11	2.57	2.90	3.22	3.97
18	0.69	1.33	1.73	2.10	2.55	2.88	3.20	3.92
19	0.69	1.33	1.73	2.09	2.54	2.86	3.17	3.88
20	0.69	1.33	1.72	2.09	2.53	2.85	3.15	3.85
21	0.69	1.32	1.72	2.08	2.52	2.83	3.14	3.82
22	0.69	1.32	1.72	2.07	2.51	2.82	3.12	3.79
23	0.69	1.32	1.71	2.07	2.50	2.81	3.10	3.77
24	0.68	1.32	1.71	2.06	2.49	2.80	3.09	3.75
25	0.68	1.32	1.71	2.06	2.49	2.79	3.08	3.73
26	0.68	1.31	1.71	2.06	2.48	2.78	3.07	3.71
27	0.68	1.31	1.70	2.05	2.47	2.77	3.06	3.69
28	0.68	1.31	1.70	2.05	2.47	2.76	3.05	3.67
29	0.68	1.31	1.70	2.05	2.46	2.76	3.04	3.66
30	0.68	1.31	1.70	2.04	2.46	2.75	3.03	3.65
40	0.68	1.30	1.68	2.02	2.42	2.70	2.97	3.55
50	0.68	1.30	1.68	2.01	2.40	2.68	2.94	3.50
60	0.68	1.30	1.67	2.00	2.39	2.66	2.91	3.46
100	0.68	1.29	1.66	1.98	2.36	2.63	2.87	3.39
∞	0.67	1.28	1.65	1.96	2.33	2.58	2.81	3.29

(9.3) Table for $\chi^2(f)$ random variable $F(x) = P(\chi^2(f) \leq x)$, where f is a parameter.

f	0.0005	0.001	0.005	0.01	0.025	0.05	0.10	0.20	0.30	0.40	0.50
1	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.06	0.15	0.27	0.45
2	0.00	0.00	0.01	0.02	0.05	0.10	0.21	0.45	0.71	1.02	1.39
3	0.02	0.02	0.07	0.11	0.22	0.35	0.58	1.01	1.42	1.87	2.37
4	0.06	0.09	0.21	0.30	0.48	0.71	1.06	1.65	2.19	2.75	3.36
5	0.16	0.21	0.41	0.55	0.83	1.15	1.61	2.34	3.00	3.66	4.35
6	0.30	0.38	0.68	0.87	1.24	1.64	2.20	3.07	3.83	4.57	5.35
7	0.48	0.60	0.99	1.24	1.69	2.17	2.73	3.82	4.67	5.49	6.35
8	0.71	0.86	1.34	1.65	2.18	2.73	3.49	4.59	5.53	6.42	7.34
9	0.97	1.15	1.73	2.09	2.70	3.33	4.17	5.38	6.39	7.36	8.34
10	1.26	1.48	2.16	2.56	3.25	3.94	4.87	6.18	7.27	8.30	9.34
11	1.59	1.83	2.60	3.05	3.82	4.57	5.58	6.99	8.15	9.24	10.34
12	1.93	2.21	3.07	3.57	4.40	5.23	6.30	7.81	9.03	10.18	11.34
13	2.31	2.62	3.57	4.11	5.01	5.89	7.04	8.63	9.93	11.13	12.34
14	2.70	3.04	4.07	4.66	5.63	6.57	7.79	9.47	10.82	12.08	13.34
15	3.11	3.48	4.60	5.23	6.26	7.26	8.55	10.31	11.72	13.03	14.34
16	3.54	3.94	5.14	5.81	6.91	7.96	9.31	11.15	12.62	13.98	15.34
17	3.98	4.42	5.70	6.41	7.56	8.67	10.09	12.00	13.53	14.94	16.34
18	4.44	4.90	6.26	7.01	8.23	9.39	10.86	12.86	14.44	15.89	17.34
19	4.91	5.41	6.84	7.63	8.91	10.12	11.65	13.72	15.35	16.85	18.34
20	5.40	5.92	7.43	8.26	9.59	10.85	12.44	14.58	16.27	17.81	19.34
21	5.90	6.45	8.03	8.90	10.28	11.59	13.24	15.44	17.18	18.77	20.34
22	6.40	6.98	8.64	9.54	10.98	12.34	14.04	16.31	18.10	19.73	21.34
23	6.92	7.53	9.26	10.20	11.69	13.09	14.85	17.19	19.02	20.69	22.34
24	7.45	8.08	9.89	10.86	12.40	13.85	15.66	18.06	19.94	21.65	23.34
25	7.99	8.65	10.52	11.52	13.12	14.61	16.47	18.94	20.87	22.62	24.34
26	8.54	9.22	11.16	12.20	13.84	15.38	17.29	19.82	21.79	23.58	25.34
27	9.09	9.80	11.81	12.88	14.57	16.15	18.11	20.70	22.72	24.54	26.34
28	9.66	10.39	12.46	13.56	15.31	16.93	18.94	21.59	23.65	25.51	27.34
29	10.23	10.99	13.12	14.26	16.05	17.71	19.77	22.48	24.58	26.48	28.34
30	10.80	11.59	13.79	14.95	16.79	18.49	20.60	23.36	25.51	27.44	29.34
40	16.91	17.92	20.71	22.16	24.43	26.51	29.05	32.34	34.87	37.13	39.34
50	23.46	24.67	27.99	29.71	32.36	34.76	37.69	41.45	46.86	51.93	59.33
60	30.34	31.74	35.53	37.48	40.48	43.19	46.46	50.64	53.81	56.62	59.33
100	59.90	61.92	67.33	70.06	74.22	77.93	82.36	87.95	92.13	95.81	99.33

Table for $\chi^2(f)$ random variable $F(x) = P(\chi^2(f) \leq x)$, where f is a parameter.

f	0.60	0.70	0.80	0.90	0.95	0.975	0.99	0.995	0.999	0.9995
1	0.71	1.07	1.64	2.71	3.84	5.02	6.63	7.88	10.83	12.12
2	1.83	2.41	3.22	4.61	5.99	7.38	9.21	10.60	13.82	15.20
3	2.95	3.66	4.64	6.25	7.81	9.35	11.34	12.84	16.27	17.73
4	4.04	4.88	5.99	7.78	9.49	11.14	13.28	14.86	18.47	20.00
5	5.13	6.06	7.29	9.24	11.07	12.83	15.09	16.75	20.52	22.11
6	6.21	7.23	8.56	10.64	12.59	14.45	16.81	18.55	22.46	24.10
7	7.28	8.38	9.80	12.02	14.07	16.01	18.48	20.28	24.32	26.02
8	8.35	9.52	11.03	13.36	15.51	17.53	20.09	21.95	26.12	27.87
9	9.41	10.66	12.24	14.68	16.92	19.02	21.67	23.59	27.88	29.67
10	10.47	11.78	13.44	15.99	18.31	20.48	23.21	25.19	29.59	31.42
11	11.53	12.90	14.63	17.28	19.68	21.92	24.72	26.76	31.26	33.14
12	12.58	14.01	15.81	18.55	21.03	23.34	26.22	28.30	32.91	34.82
13	13.64	15.12	16.98	19.81	22.36	24.74	27.69	29.82	34.53	36.48
14	14.69	16.22	18.15	21.06	23.68	26.12	29.14	31.32	36.12	38.11
15	15.73	17.32	19.31	22.31	25.00	27.49	30.58	32.80	37.70	39.72
16	16.78	18.42	20.47	23.54	26.30	28.85	32.00	34.27	39.25	41.31
17	17.82	19.51	21.61	24.77	27.59	30.19	33.41	35.72	40.79	42.88
18	18.87	20.60	22.76	25.99	28.87	31.53	34.81	37.16	42.31	44.43
19	19.91	21.69	23.90	27.20	30.14	32.85	36.19	38.58	43.82	45.97
20	20.95	22.77	25.04	28.41	31.41	34.17	37.57	40.00	45.31	47.50
21	21.99	23.86	26.17	29.62	32.67	35.48	38.93	41.40	46.80	49.01
22	23.03	24.94	27.30	30.81	33.92	36.78	40.29	42.80	48.27	50.51
23	24.07	26.02	28.43	32.01	35.17	38.08	41.64	44.18	49.73	52.00
24	25.11	27.10	29.55	33.20	36.42	39.36	42.98	45.56	51.18	53.48
25	26.14	28.17	30.68	34.38	37.65	40.65	44.31	46.93	52.62	54.95
26	27.18	29.25	31.79	35.56	38.89	41.92	45.64	48.29	54.05	56.41
27	28.21	30.32	32.91	36.74	40.11	43.19	46.96	49.64	55.48	57.86
28	29.25	31.39	34.03	37.92	41.34	44.46	48.28	50.99	56.89	59.30
29	30.28	32.46	35.14	39.09	42.56	45.72	49.59	52.34	58.30	60.73
30	31.32	33.53	36.25	40.26	43.77	46.98	50.89	53.67	59.70	62.16
40	41.62	44.16	47.27	51.81	55.76	59.34	63.69	66.77	73.40	76.09
50	51.89	54.72	58.16	63.17	67.50	71.42	79.49	79.49	86.66	89.56
60	62.13	65.23	68.97	74.40	79.08	83.30	88.38	91.95	99.61	102.69
100	102.95	106.91	111.67	118.50	124.34	129.56	135.81	140.17	149.45	153.17

(9.4) Table for Binomial random variable $P(Bin(n, p) \leq k)$ if $p \leq 0.5$.
 If $p > 0.5$, then $P(Bin(n, p) \leq k) = P(Bin(n, 1 - p) \geq n - k)$.

n	k	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500
	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975	0.7500
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250
	1	0.9928	0.9720	0.9392	0.8960	0.8438	0.7840	0.7183	0.6480	0.5747	0.5000
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1256	0.0915	0.0625
	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4735	0.3910	0.3125
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0313
	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562	0.1875
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156
	1	0.9672	0.8847	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333	0.1636	0.1094
7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078
	1	0.9566	0.8503	0.7166	0.5767	0.4449	0.3294	0.2338	0.1586	0.1024	0.0625
8	0	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0319	0.0168	0.0084	0.0039
	1	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1691	0.1064	0.0632	0.0352
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020
	1	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.1211	0.0705	0.0385	0.0195

Table for Binomial random variable $P(Bin(n, p) \leq k)$ if $p \leq 0.5$.
 If $p > 0.5$, then $P(Bin(n, p) \leq k) = P(Bin(n, 1 - p) \geq n - k)$.

n	k	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107
11	0	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0088	0.0036	0.0014	0.0005
	1	0.8981	0.6974	0.4922	0.3221	0.1971	0.1130	0.0606	0.0302	0.0139	0.0059
12	0	0.5404	0.2824	0.1422	0.0687	0.0317	0.0138	0.0057	0.0022	0.0008	0.0002
	1	0.8816	0.6590	0.4435	0.2749	0.1584	0.0850	0.0424	0.0196	0.0083	0.0032

Table for Poisson random variable $P(Po(\mu) \leq k)$.

k	μ																
	9.0	9.5	10.0	11.0	12.0	13.0	14.0	15.0	16.0	17.0							
0	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000							
1	0.0012	0.0008	0.0005	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000							
2	0.0062	0.0042	0.0028	0.0012	0.0005	0.0002	0.0001	0.0000	0.0000	0.0000							
3	0.0212	0.0149	0.0103	0.0049	0.0023	0.0011	0.0005	0.0002	0.0001	0.0000							
4	0.0550	0.0403	0.0293	0.0151	0.0076	0.0037	0.0018	0.0009	0.0004	0.0002							
5	0.1157	0.0885	0.0671	0.0375	0.0203	0.0107	0.0055	0.0028	0.0014	0.0007							
6	0.2068	0.1649	0.1301	0.0786	0.0458	0.0259	0.0142	0.0076	0.0040	0.0021							
7	0.3239	0.2687	0.2202	0.1432	0.0895	0.0540	0.0316	0.0180	0.0100	0.0054							
8	0.4557	0.3918	0.3328	0.2320	0.1550	0.0998	0.0621	0.0374	0.0220	0.0126							
9	0.5874	0.5218	0.4579	0.3405	0.2424	0.1658	0.1094	0.0699	0.0433	0.0261							
10	0.7060	0.6453	0.5830	0.4599	0.3472	0.2517	0.1757	0.1185	0.0774	0.0491							
11	0.8030	0.7520	0.6968	0.5793	0.4616	0.3532	0.2600	0.1848	0.1270	0.0847							
12	0.8758	0.8364	0.7916	0.6887	0.5760	0.4631	0.3585	0.2676	0.1931	0.1350							
13	0.9261	0.8981	0.8645	0.7813	0.6815	0.5730	0.4644	0.3632	0.2745	0.2009							
14	0.9585	0.9400	0.9165	0.8540	0.7720	0.6751	0.5704	0.4657	0.3675	0.2808							
15	0.9780	0.9665	0.9513	0.9074	0.8444	0.7636	0.6694	0.5681	0.4667	0.3715							
16	0.9889	0.9823	0.9730	0.9441	0.8987	0.8355	0.7559	0.6641	0.5660	0.4677							
17	0.9947	0.9911	0.9857	0.9678	0.9370	0.8905	0.8272	0.7489	0.6593	0.5640							
18	0.9976	0.9957	0.9928	0.9823	0.9626	0.9302	0.8826	0.8195	0.7423	0.6550							
19	0.9989	0.9980	0.9965	0.9907	0.9787	0.9573	0.9235	0.8752	0.8122	0.7363							
20	0.9996	0.9991	0.9984	0.9953	0.9884	0.9750	0.9521	0.9170	0.8682	0.8055							
21	0.9998	0.9996	0.9993	0.9977	0.9939	0.9859	0.9712	0.9469	0.9108	0.8615							
22	0.9999	0.9999	0.9997	0.9990	0.9970	0.9924	0.9833	0.9673	0.9418	0.9047							
23	1.0000	0.9999	0.9999	0.9995	0.9985	0.9960	0.9907	0.9805	0.9633	0.9367							
24	1.0000	1.0000	1.0000	0.9998	0.9993	0.9980	0.9950	0.9888	0.9777	0.9594							
25	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9974	0.9938	0.9869	0.9748							
26	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9987	0.9967	0.9925	0.9848							
27	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994	0.9983	0.9959	0.9912							
28	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9991	0.9978	0.9950							
29	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9994	0.9986							
30	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9997	0.9993							
31	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9996							
32	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9996							
33	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9999							
34	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999							
35	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000							