

PLEASE NOTE THAT

(a) I did not set this exam, so the style is different from the Fall 2018 version

(b) The information provided at the end does NOT match what will be provided in the Fall 2018 version.

(c) We have covered all the material in this year's course to enable you to answer all questions APART FROM ONE PART - I will leave it to you to spot which part this is. Other than that, the level and content of the questions is appropriate.

(d) I DO NOT HAVE SOLUTIONS TO THIS EXAM - if that worries you, turn back now and forget you ever saw it. DAS, 28/11

McGill University
Faculty of Science

Winter 2016
Final Examination

Math 323: Probability

Tuesday April 19, 2016

Time: 2:00 PM - 5:00 PM

Sketch
Solutions

Examiner:

Associate Examiner:

Student name (last, first)	Student number (McGill ID)

INSTRUCTIONS

This exam contains 17 pages (including this cover page) and 8 problems. Formula sheets and selected theorems and definitions are provided on the last three pages. The following rules apply:

- You may **not** use your textbook or notes during this exam.
- **You are permitted a scientific calculator.** No graphing calculators, computers, etc.
- **Organize your work**, in a reasonably neat and coherent way, in the space provided.
- **Mysterious or unsupported answers will not receive full credit.** A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit.
- If you need more space, use the back of the pages; clearly indicate when you have done this.

Problem	Points	Score
1	16	
2	10	
3	7	
4	9	
5	9	
6	14	
7	16	
8	9	
Total:	90	

1. Suppose that a drawer contains 21 loose socks; seven black socks, eight white socks and six green socks. Two socks are drawn randomly (without replacement) from the drawer. Let X be the number of black socks picked and Y be the number of white socks picked.

- (a) (3 points) What is the probability that the two socks chosen from the drawer are the same colour?

Count of $2B: \binom{7}{2}$ $2W: \binom{8}{2}$ $2G: \binom{6}{2}$ Total #: $\binom{21}{2}$

$$\therefore P(\text{Same colour}) = P(2B) + P(2W) + P(2G) = \frac{\binom{7}{2} + \binom{8}{2} + \binom{6}{2}}{\binom{21}{2}}$$

$$= \frac{21 + 28 + 15}{210} = \frac{32}{105}$$

- (b) (6 points) Find the joint probability mass function of X and Y . Summarize these probabilities in a table rather than expressing them in functional form.

Hypergeometric/Partitioning gives the general formula

$$\binom{7}{x} \binom{8}{y} \binom{6}{2-x-y} \quad \text{for } \begin{cases} 0 \leq x \leq 7 \\ 0 \leq y \leq 8 \\ 0 \leq 2-x-y \leq 6 \end{cases}$$

i.e. need $0 \leq x \leq 2$, $0 \leq y \leq 2$ and $0 \leq x+y \leq 2$.

$p(x,y) \times 210$

		Y		
		0	1	2
X	0	15	48	28
	1	42	56	0
	2	21	0	0

$$\text{eg } P(X=1, Y=0) = \frac{\binom{7}{1} \binom{8}{0} \binom{6}{1}}{\binom{21}{2}}$$

$$= \frac{7 \times 6}{210} = \frac{42}{210}$$

etc.

(c) (4 points) Compute $\text{Cov}(X, Y)$

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

$$E[XY] = \sum_{x=0}^2 \sum_{y=0}^2 xy p(x, y) = p(1, 1) = \frac{56}{210}.$$

$$E[X] = \sum_{x=0}^2 x p_X(x) = 1 \cdot p_X(1) + 2 \cdot p_X(2) = \left(\frac{42+56}{210}\right) + \left(\frac{2 \times 21}{210}\right) = \frac{140}{210}$$

$$E[Y] = \sum_{y=0}^2 y p_Y(y) = 1 \cdot p_Y(1) + 2 \cdot p_Y(2) = \left(\frac{42+56}{210}\right) + \left(\frac{2 \times 28}{210}\right) = \frac{160}{210}.$$

$$\therefore \text{Cov}(X, Y) = \frac{56}{210} - \frac{(140 \times 160)}{210^2} \approx -0.2413.$$

(d) (3 points) Given that the selected socks match, what is the conditional probability that they were green?

$$P(2G | \text{Match}) = \frac{P(2G \cap \text{Match})}{P(\text{Match})} = \frac{\binom{6}{2}}{\binom{7}{2} + \binom{8}{2} + \binom{6}{2}} = \frac{15}{64}.$$

2. Jill's bowling scores are approximately normally distributed with mean 170 and standard deviation 20, while Jack's scores are approximately normally distributed with mean 160 and standard deviation 15. If Jack and Jill each bowl one game, then assuming that their scores are independent random variables, approximate the probability that

(a) (3 points) Jill scores above 135,

$$\text{Jack : } X \sim N(160, 15^2) \quad \text{Jill : } Y \sim N(170, 20^2) \quad Z \sim N(0, 1)$$

$$\begin{aligned} P(Y > 135) &= P\left(\frac{Y-170}{20} > \frac{135-170}{20}\right) = P\left(Z > \frac{-7}{4}\right) = 1 - P\left(Z > \frac{7}{4}\right) \\ &= 1 - P(Z > 1.75) \\ &= 1 - 0.0401 \\ &= 0.9599 \end{aligned}$$

(b) (4 points) Jack's score is higher than Jill's score.

$$P(X > Y) = P(X - Y > 0)$$

But by transformation results (using mgf) $X - Y \sim N(-10, 15^2 + 20^2)$.
 $\equiv N(-10, 25^2)$.

$$\begin{aligned} \therefore P(X - Y > 0) &= P\left(\frac{(X - Y) - (-10)}{25} > \frac{-10}{25}\right) = P\left(Z > -\frac{2}{5}\right) = P\left(Z > \frac{2}{5}\right) \\ &= 1 - P(Z > 0.4) \\ &= 1 - 0.3446 = \end{aligned}$$

(c) (3 points) If X and Y are the normal RVs that approximate Jack and Jill's respective bowling scores, find $E[(X - Y)^2]$.

$$\begin{aligned} \text{Let } W &= (X - Y). \quad E[W^2] = \text{Var}[W] + \{E[W]\}^2. \quad (\text{as } \text{Var}[W] = E[W^2] - \{E[W]\}^2) \\ &= (25)^2 + (-10)^2 \\ &= 725. \end{aligned}$$

3. (7 points) Let A and B be independent events. Prove algebraically (without Venn Diagrams) that $\bar{A}$ and $\bar{B}$ are independent of each other.

$$\begin{aligned}P(\bar{A} \cap \bar{B}) &= P(\overline{A \cup B}) && \text{(de Morgan)} \\&= 1 - P(A \cup B) && \text{(complementation)} \\&= 1 - (P(A) + P(B) - P(A \cap B)) && \text{(general addition formula)} \\&= 1 - P(A) - P(B) + P(A \cap B) \\&= 1 - P(A) - P(B) + P(A)P(B) && \text{(independence)} \\&= (1 - P(A))(1 - P(B)) \\&= P(\bar{A})P(\bar{B})\end{aligned}$$

$\therefore \bar{A}$ and $\bar{B}$ are independent.

4. A random variable X following a *Laplace distribution* has the following pdf

$$f(x) = \frac{1}{2} \lambda e^{-\lambda|x|} \quad \text{for } -\infty < x < \infty,$$

where the parameter $\lambda > 0$.

(a) (5 points) Show that the moment generating function of X is given by $\frac{\lambda^2}{\lambda^2 - t^2}$ when $|t| < \lambda$. For suitable t

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{2} \lambda e^{-\lambda|x|} dx.$$

$$= \int_{-\infty}^0 e^{tx} \frac{1}{2} \lambda e^{-\lambda|x|} dx + \int_0^{\infty} e^{tx} \frac{1}{2} \lambda e^{-\lambda|x|} dx$$

$$= \int_{-\infty}^0 e^{tx} \frac{1}{2} \lambda e^{-\lambda(-x)} dx + \int_0^{\infty} e^{tx} \frac{1}{2} \lambda e^{-\lambda x} dx$$

$$= \int_0^{\infty} e^{-tx} \frac{1}{2} \lambda e^{-\lambda x} dx + \int_0^{\infty} e^{tx} \frac{1}{2} \lambda e^{-\lambda x} dx.$$

$$= \frac{1}{2} \lambda \left[\int_0^{\infty} e^{-(t+\lambda)x} dx + \int_0^{\infty} e^{-(\lambda-t)x} dx \right]$$

$$= \frac{1}{2} \lambda \left[\left[\frac{-1}{(t+\lambda)} e^{-(t+\lambda)x} \right]_0^{\infty} + \left[\frac{-1}{(\lambda-t)} e^{-(\lambda-t)x} \right]_0^{\infty} \right] \text{ if } |t| < \lambda.$$

$$= \frac{1}{2} \lambda \left[\frac{1}{\lambda+t} + \frac{1}{\lambda-t} \right] = \frac{1}{2} \lambda \left[\frac{2\lambda}{\lambda^2 - t^2} \right]$$

$$= \frac{\lambda^2}{\lambda^2 - t^2} \quad |t| < \lambda$$

- (b) (4 points) Use the MGF from (a) to show that the mean and variance of the Laplace distribution are 0 and $2/\lambda^2$, respectively.

We have

$$\begin{aligned} E[X] &= \left. \frac{d}{dt} \{M_X(t)\} \right|_{t=0} = m_X^{(1)}(0) = -\lambda^2 (\lambda^2 - t^2)^{-2} \cdot 2t \Big|_{t=0} \\ &= 0. \end{aligned}$$

$$\begin{aligned} E[X^2] &= \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = \left. \frac{d}{dt} \left\{ \frac{2\lambda^2 t}{(\lambda^2 - t^2)^2} \right\} \right|_{t=0} \\ &= 2\lambda^2 \left[\frac{(\lambda^2 - t^2)^2 - t(\lambda^2 - t^2) \cdot (-2t)}{(\lambda^2 - t^2)^4} \right]_{t=0} \\ &= \frac{2\lambda^2}{\lambda^4} = \frac{2}{\lambda^2}. \end{aligned}$$

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

Note also that

$$E[X] = \int_{-\infty}^{\infty} x \cdot \frac{1}{2} \lambda e^{-\lambda|x|} dx = 0.$$

as integrand is an integrable odd function around zero.

5. Consider the following joint probability density function of X and Y

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x+y)} & \text{for } x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) (6 points) Find the probability density function of $Z = \frac{X+Y}{2}$.

Note that $f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \forall x,y \quad \therefore X, Y \text{ independent}$
with $X \sim \text{Exponential}(1)$ and $Y \sim \text{Exponential}(1)$.

$\therefore$ By mgsf

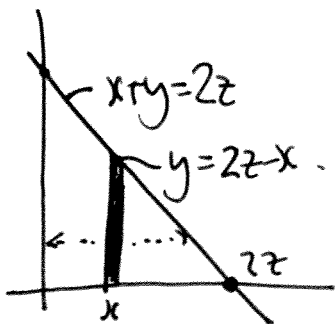
$$M_Z(t) = M_X\left(\frac{t}{2}\right)M_Y\left(\frac{t}{2}\right) = \left(\frac{1}{1-t/2}\right)\left(\frac{1}{1-t/2}\right) = \left(\frac{1}{1-t/2}\right)^2$$

$\therefore Z \sim \text{Gamma}(2, 1/2)$ (from formula sheet)

with $E[Z] = 1$.

$$\therefore f_Z(z) = \frac{z^2}{\Gamma(2)} z^{2-1} e^{-2z} = 4ze^{-2z} \quad z > 0, \quad z = 0 \text{ o.w.}$$

Alternatively: $F_Z(z) = P(Z \leq z) = P\left(\frac{X+Y}{2} \leq z\right) = P(X+Y \leq 2z) \quad z > 0$.



$$= \int_0^{2z} \left\{ \int_0^{2z-x} f_{X,Y}(x,y) dy \right\} dx.$$

$$= \int_0^{2z} \left\{ \int_0^{2z-x} e^{-(x+y)} dy \right\} dx = \int_0^{2z} e^{-x} \left\{ \int_0^{2z-x} e^{-y} dy \right\} dx$$

$$= \int_0^{2z} \left\{ \left[-e^{-y} \right]_0^{2z-x} \right\} e^{-x} dx = \int_0^{2z} (1 - e^{-(2z-x)}) e^{-x} dx.$$

$$= \left[-e^{-x} \right]_0^{2z} - \int_0^{2z} e^{-2z} dx = 1 - e^{-2z} - 2ze^{-2z} \quad z > 0.$$

$$\therefore f_Z(z) = 4ze^{-2z} \quad z > 0 \text{ as before.}$$

(cont'd) Consider the following joint probability density function of X and Y

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x+y)} & \text{for } x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

(b) (3 points) Are X and Y independent? (Justify your answer)

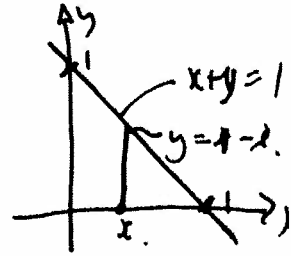
Yes: $f_{X,Y}(x,y) = f_X(x)f_Y(y) \quad \forall (x,y) \in \mathbb{R}^2$.

$$f_X(x) = e^{-x} \quad x \geq 0, \text{ zero o.w.}$$

$$f_Y(y) = e^{-y} \quad y \geq 0, \text{ zero o.w.}$$

6. Suppose X and Y have the joint pdf given below,

$$f_{X,Y}(x,y) = \begin{cases} 24xy & \text{for } 0 < x < 1, 0 < y < 1, x+y < 1 \\ 0 & \text{otherwise} \end{cases}$$



(a) (3 points) Find the marginal pdf of X (remember to specify the support).

For $0 < x < 1$.

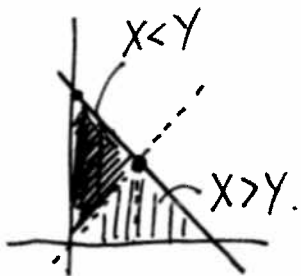
$$\begin{aligned} f_X(x) &= \int_0^{1-x} f_{X,Y}(x,y) dy = \int_0^{1-x} 24xy dy = 24x \left[\frac{y^2}{2} \right]_0^{1-x} \\ &= 12x(1-x)^2 \quad 0 < x < 1 \quad \text{zero a.s.} \\ &\Rightarrow X \sim \text{Beta}(2,3). \end{aligned}$$

(b) (5 points) Find $P(Y > X)$.

As X and Y are identically distributed / $f_{X,Y}(x,y)$ is symmetric in x and y , with a 'symmetric' support, we must have $P(X > Y) = 1/2 = P(Y > X)$.

Explicitly:

$$P(Y > X) = \int_0^{1/2} \left\{ \int_x^{1-x} 24xy dy \right\} dx = \int_0^{1/2} 24x \left[\frac{y^2}{2} \right]_x^{1-x} dx.$$



$$\begin{aligned} &= \int_0^{1/2} 12x[(1-x)^2 - x^2] dx = \int_0^{1/2} 12x(1-2x) dx \\ &= \left[6x^2 - 8x^3 \right]_0^{1/2} = \frac{6}{4} - 1 = \frac{1}{2}. \end{aligned}$$

(cont'd) Suppose X and Y have the joint pdf given below,

$$f_{X,Y}(x,y) = \begin{cases} 24xy & \text{for } 0 < x < 1, 0 < y < 1-x \\ 0 & \text{otherwise} \end{cases}$$

(c) (3 points) Find $E[XY]$.

$$\begin{aligned} E[XY] &= \int_0^1 \left[\int_0^{1-x} xy \cdot (24xy) dy \right] dx = 24 \int_0^1 \left[\int_0^{1-x} x^2 y^2 dy \right] dx \\ &= 24 \int_0^1 \left[\frac{y^3}{3} \right]_0^{1-x} x^2 dx \\ &= 8 \int_0^1 x^2 (1-x)^3 dx \quad \text{- integrand is prop' to } \text{Beta}(3,4) \\ &= 8 \cdot \frac{\Gamma(3)\Gamma(4)}{\Gamma(7)} = 8 \cdot \frac{2!3!}{6!} = \frac{8}{10}. \end{aligned}$$

(d) (3 points) Find the conditional pdf of Y given that $X = 1/3$ (remember to specify the support).

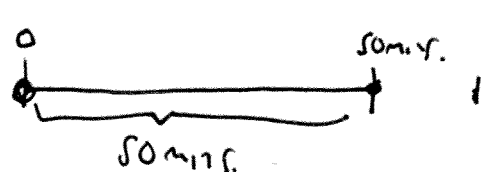
$$\begin{aligned} f_{Y|X}(y|x) &= \frac{f_{X,Y}(x,y)}{f_X(x)} \quad 0 < y < 1-x \quad \text{for } 0 < x < 1 \\ &= \frac{24xy}{12x(1-x)^2} = \frac{2}{(1-x)^2} y \quad 0 < y < 1-x \\ &\quad \text{zero o.w.} \end{aligned}$$

$$\text{At } x = 1/3$$

$$f_{Y|X}(y|1/3) = \frac{2}{(2/3)^2} y = \frac{9}{2} y \quad 0 < y < \frac{2}{3} \quad \text{zero o.w.}$$

7. The number of planes arriving at a small private airport in a day is a random variable having a Poisson distribution with a constant rate (day and night) of $\lambda = 28.8$ per day.

- (a) (3 points) What is the probability that the time between two consecutive arrivals is at least 50 minutes?

$$X \sim \text{Poisson}(\lambda) \quad \lambda = 28.8 \text{ per day : constant rate across time.}$$


$$P(\text{No planes in 50 mins}) = P(Y=0) \quad Y \sim \text{Poisson}\left(\frac{28.8}{24 \times 60} \times 50\right)$$

$$= \exp\{-\lambda_Y\} \quad = \lambda_Y \text{ say}$$

- (b) (6 points) Monica begins her eight hour shift at the airport at 8am. This shift is divided into four two-hour periods. What is the probability that at least two planes arrive during exactly three of these four periods?

$$P(\text{Two planes in 8 hrs at least}) = 1 - P(\text{zero planes}) - P(\text{one plane}).$$

$$\text{Let } Z = \text{"Number of planes per shift"} \quad Z \sim \text{Poisson}\left(28.8 \times \frac{8}{24}\right)$$

$$\text{Then } P(Z=z) = \frac{e^{-\lambda_Z} \lambda_Z^z}{z!} \quad = \lambda_Z \text{ say.}$$

$$\therefore P(\geq 2 \text{ planes}) = 1 - e^{-\lambda_Z} - \lambda_Z e^{-\lambda_Z} = p_2 \text{ say.}$$

Over 4 shifts: need Binomial distⁿ.

$$\binom{4}{3} p_2^3 (1-p_2)^1$$

(Shift counts are independent by Poisson definition as limit of iid Binomial).

- (c) (2 points) If we define a random variable X as the waiting time from now until the arrival of the second plane, what distribution does X follow? (Be sure to specify the values of the parameters of the distribution)

$$X \sim \text{Gamma}(2, 1/\lambda) \quad \lambda = 28.8$$

[We did not cover the Poisson process]. (X in hours).

- (d) (5 points) In reference to part (c), find the probability that X is greater than 5 hours.

$$P(X > 5) = \int_5^{\infty} f_X(x) dx$$

$$= \int_5^{\infty} \frac{\lambda^2}{\Gamma(2)} x e^{-\lambda x} dx.$$

$$= \lambda^2 \left[-\frac{1}{\lambda} x e^{-\lambda x} \right]_5^{\infty} + \lambda \int_5^{\infty} e^{-\lambda x} dx.$$

$$= 5\lambda e^{-5\lambda} + \lambda \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_5^{\infty}$$

$$= 5\lambda e^{-5\lambda} + e^{-5\lambda} = (5\lambda + 1)e^{-5\lambda}$$

8. The annual proportion of erroneous income tax returns filed to the Canada Revenue Agency is a random variable having a beta distribution with $\alpha = 2$ and $\beta = 8$.

(a) (5 points) What is the probability that there will be fewer than 10 percent erroneous returns in a given year?

$$X \sim \text{Beta}(2, 8) \quad f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad 0 < x < 1$$

$$= \frac{\Gamma(10)}{\Gamma(2)\Gamma(8)} x(1-x)^7 = \frac{9!}{7!} x(1-x)^7 = 72x(1-x)^7$$

$$P(X < 0.10) = \int_0^{0.1} 72x(1-x)^7 dx = 72 \left[-\frac{x(1-x)^8}{8} \right]_0^{0.1} + 72 \int_0^{0.1} \frac{1}{8} (1-x)^8 dx$$

$$= \frac{72}{8} \left\{ -0.1(0.9)^8 \right\} + \left[-(1-x)^9 \right]_0^{0.1} = 1 - (0.9)^9 - (0.9)^9 = 1 - 2(0.9)^9$$

$$= 0.2251$$

- (b) (4 points) Assuming there are no erroneous returns, the cost of processing all tax returns is 3.8 million. With each percent increase of erroneous returns the processing cost increases by \$42,800. Compute the expected yearly total cost, and the corresponding standard deviation. Recall that the mean and variance of the beta distribution are given by

$$\mu = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Let T be the total cost. Then

$$T = 3.8 \times 10^6 + 4.28 \times 10^4 X \quad \text{— compute numerically}$$

$$= a + bX \quad \text{say.}$$

$$\text{Then } E[T] = a + bE[X] = a + b \frac{8}{10}$$

$$\text{Var}[T] = b^2 \text{Var}[X] = b^2 \left(\frac{2 \times 8}{10^2 \cdot 11} \right) = b^2 \frac{16}{1100}.$$

Special probability distributions

1. Bernoulli distribution, $f(x; p) = p^x(1-p)^{1-x}$, for $x = 0, 1$.
2. Binomial distribution, $f(x; n, p) = \binom{n}{x} p^x(1-p)^{n-x}$, for $x = 0, 1, \dots, n$.
3. Discrete uniform distribution, $f(x; k) = \frac{1}{k}$, for $x = 1, 2, \dots, k$.
4. Geometric distribution, $f(x; p) = p(1-p)^{x-1}$, for $x = 1, 2, \dots$.
6. Negative binomial, $f(x; k, p) = \binom{x-1}{k-1} p^k(1-p)^{x-k}$ for $x = k, k+1, k+2, \dots$.
7. Poisson distribution, $f(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$, for $x = 0, 1, 2, 3, \dots$.
8. Hypergeometric Distribution, $f(x_1, \dots, x_k; n, M_1, \dots, M_k) = \frac{\binom{M_1}{x_1} \binom{M_2}{x_2} \dots \binom{M_k}{x_k}}{\binom{N}{n}}$,
for $x_i = 0, 1, \dots, n$ and $x_i \leq M_i$ for each i , where $\sum_{i=1}^k x_i = n$, and $\sum_{i=1}^k M_i = N$.

Special probability densities

1. Beta distribution, $f(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$ for $0 < x < 1$, where $\alpha, \beta > 0$.
2. Exponential distribution, $f(x; \lambda) = \lambda e^{-\lambda x}$, for $x > 0, \lambda > 0$.
3. Gamma distribution, $f(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$ for $x > 0$, where $\alpha > 0$ and $\beta > 0$.
4. Normal distribution, $f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, for $-\infty < x < \infty$, where $\sigma > 0$.
5. Uniform distribution. $f(x; \alpha, \beta) = \frac{1}{\beta - \alpha}$, for $\alpha < x < \beta$.
6. Student- t distribution $f(t; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$ for $-\infty < t < \infty$.

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

Theorem 2 If a and b are constants, then $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$.

Theorem 3 If X has variance σ_x^2 , then the variance of $U = aX + c$ is $\text{Var}(U) = a^2 \text{Var}(X) = a^2 \sigma_x^2$.

MGF The moment generating function of a random variable X is defined by $M_X(t) = \mathbb{E}[e^{tX}]$

Theorem 4 $M_{\frac{X+a}{b}}(t) = \mathbb{E}[e^{(\frac{X+a}{b})t}] = e^{\frac{a}{b}t} M_X(\frac{t}{b})$

Theorem 5 For two independent RVs, we have $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$

Binomial Theorem For any positive integer n : $(x + y)^n = \sum_{r=0}^n \binom{n}{r} x^{n-r} y^r$.

Geometric Series $\sum_{k=1}^{\infty} ar^k = \frac{ar}{(1-r)}$ or equivalently $\sum_{k=0}^{\infty} ar^k = \frac{a}{(1-r)}$ if $-1 < |r| < 1$

Bayes Theorem If $B_1, B_2, \dots, B_k$ constitute a partition of the sample space S and $P(B_i) \neq 0$ for $i = 1, 2, \dots, k$, then for any event A in S such that $P(A) \neq 0$, $P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^k P(A|B_i)P(B_i)}$ for $j = 1, 2, \dots, k$.

Chebyshev's Theorem If μ and σ are the mean and the standard deviation of a random variable X , then for any positive constant k , we have $P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$.

Theorem 6 $\sigma_{XY} = \text{Cov}(X, Y) \equiv \mathbb{E}[(X - \mu_X) \cdot (Y - \mu_Y)] = \mathbb{E}[XY] - \mu_X \cdot \mu_Y$

Theorem 5 $\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X, Y)$

Theorem 6 If X and Y are independent, then $E[XY] = E[X] \cdot E[Y]$ and $\sigma_{X,Y} = 0$.

Theorem 7 If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$.

Theorem 8 If $X_1, X_2, \dots, X_n$ are mutually independent normal random variables with means $\mu_1, \mu_2, \dots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, then the linear combination: $Y = \sum_{i=1}^n c_i X_i$ follows the normal distribution: $N(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n c_i^2 \sigma_i^2)$.

Theorem 9 If $X_1, \dots, X_n$ constitute a random sample from an infinite population with the mean μ and the variance σ^2 , then, $\mathbb{E}(\bar{X}) = \mu$ and $\text{Var}(\bar{X}) = \sigma^2/n$.

Theorem 10 If $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$ and $\bar{X}$ is the sample mean, then $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

Central Limit Theorem If $X_1, \dots, X_n$ constitute a random sample from an infinite population with the mean μ and the variance σ^2 , then $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ is approximately distributed according to $N(0, 1)$ for large n .

Theorem 11 If $X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} N(0, 1)$, then $Y = \sum_{i=1}^n X_i^2 \sim \chi_n^2$

Theorem 12 If Y and Z are independent RVs, where $Y \sim \chi_\nu^2$ and $Z \sim N(0, 1)$, then a random variable $T = \frac{Z}{\sqrt{Y/\nu}}$ follows a t -distribution with ν degrees of freedom, denoted by $T \sim t_\nu$.

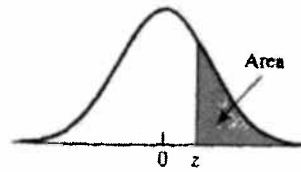
Theorem 13 If $\bar{X}$ and S^2 are sample mean and sample variance of a random sample of size n from a normal population with a mean μ and the variance σ^2 , then the t -statistic $T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$.

Theorem 14 If U and V are independent RVs with $U \sim \chi_{\nu_1}^2$ and $V \sim \chi_{\nu_2}^2$, then $F = \frac{U/\nu_1}{V/\nu_2}$ is a random variable having an F distribution with ν_1 and ν_2 degrees of freedom denoted by F_{ν_1, ν_2} .

Theorem 15 If S_1^2 and S_2^2 are the sample variances of two independent random samples of size n_1 and n_2 from normal populations with the variances σ_1^2 and σ_2^2 , then $F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} \sim F_{n_1-1, n_2-1}$ and F is called an F -statistic.

Normal Curve Areas

Standard normal probability in right-hand tail (for negative values of z , areas are found by symmetry)



Second decimal place of z										
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000 233									
4.0	.000 031 7									
4.5	.000 003 40									
5.0	.000 000 287									

From R. E. Walpole, *Introduction to Statistics* (New York: Macmillan, 1968).