The purpose of the scoring guide is to provide a uniform evaluation of the tests by the graders. Therefore, the guide gives you the main ideas (and at least one possible solution) for each problem and the associated scores. This guide is not intended to be a detailed description of a complete solution to each problem; the steps described are sketches of a solution that achieves maximum score.
The partial scores given in the guide are only awarded to the solver if the related idea is included in the test as a step in a clearly described and reasoned solution. Thus, for example, the mere description of statements or definitions in the material without its application (even if any of the facts described do indeed play a role in the solution) is not worth the point.
A partial score is given for any idea or part solution from which the correct solution of the problem could be obtained with the appropriate addition of the reasoning described in the test. If a solver starts several solutions that are significantly different, then you can give score for at most one. If each described solution or part of solution is correct or can be extended to be correct, then the solution with the most points will be evaluated. However, if there are correct and incorrect solutions between several solution attempts and it is not clear from the test which one the solver considered to be correct, then the solution with fewer points will be evaluated (even if this score is 0). The scores in this guide may be further subdivided as necessary. Of course, a good solution other than that described in this guide is also worth the maximum score.
In the event of an arithmetic error, one point is deducted from each exercise. An exception to this is when computing the numbers significantly simplifies or modifies the structure of the task. In these cases, no points are awarded to the steps which follow the computation and are missing from the solution of the student.

---

1. State the following definitions/theorems.
   (a) What is the Moment generating function $M_X(t)$ of a random variable X?
   (b) How are the moment generating functions of $aX, X+b, X+Y$ related to the moment generating functions of X and Y for some constants $a, b \in \mathbb{R}$ and given that X and Y are independent? *Solution:*

   (5 points) For any random variable X, the moment generating function $M_X(t) = \mathbb{E}(e^{Xt})$, provided that this expected value exists. It is called the moment generating function because $M_X^{(i)}(0) = \mathbb{E}(X^i)$, and can be used to find all moments of the random variable X.
   (5 points) $M_{aX}(t) = M_X(at)$
   (5 points) $M_{X+b}(t) = e^{bt}M_X(t)$
   (5 points) $M_{X+Y}(t) = M_X(t)M_Y(t)$ when X, Y are independent.

2. A fair coin is tossed 3 times. Let X denote the number of Heads, and let Y denote the absolute value of the difference between the number of Heads and the number of Tails (for example, if there is 1 Head and 2 Tails, then $Y = 1$).
   (a) What is the joint probability mass function of X and Y?
   (b) What is $Cov(X, Y)$? (c) Are X and Y independent? *Solution:*

   (5 points)

$Y \backslash X$	0	1	2	3	p_Y
1	0	$\frac{3}{8}$	$\frac{3}{8}$	0	$\frac{6}{8}$
3	$\frac{1}{8}$	0	0	$\frac{1}{8}$	$\frac{2}{8}$
p_X	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

   (3 points) We know that $\mathbb{E}(g(X,Y)) = \sum_x \sum_y g(x,y)p_{X,Y}(x,y)$.

   (2 points) So, $\mathbb{E}(X) = \frac{3}{2}$

(2 points) $\mathbb{E}(Y) = \frac{3}{2}$

(3 points) $\mathbb{E}(XY) = \frac{9}{4}$

(2 points) so, $Cov(X, Y) = 0$
(3 points) They are not independent because $p_{X,Y}(x, y) \neq p_X(x)p_Y(y)$ for all values of x, y. In particular, $p_{X,Y}(0, 1) = 0 \neq p_X(0)p_Y(1) = \frac{3}{32}$.

3. Let X and Y be two random variables with the following joint probability density function:

$$f_{X,Y}(x, y) = \begin{cases} \alpha xy & \text{if } 0 < x < y < 1, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Determine α. (b) Find $f_{X|Y}(x|y)$ and $E(X|Y)$. *Solution:*

(2 points) Since $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)dxdy = 1$
(2 points) good picture of the range of X, Y with the line of integral.
(2 points) so, $\int_0^1 \int_0^y \alpha xydxdy = 1$
(2 points) solving which we get, $\alpha = 8$.
(2 points) Since $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dx$,
(2 points) we will get $f_Y(y) = \int_0^y 8xydx = 4y^3$,
(2 points) We also know that $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$ when $f_Y(y) \neq 0$, so we get,
(2 points) $f_{X|Y}(x|y) = \frac{8xy}{4y^3} = \frac{2x}{y^2}$ when $0 < x < y < 1$.
(2 points) We also know that $E(X|Y) = \int_{-\infty}^{\infty} xf_{X|Y}(x|y)dx$,
(2 points) so, $E(X|Y) = \int_0^y x\frac{2x}{y^2}dx = \frac{2y}{3}$

4. The life of bubbles (time until they burst), generated by a bubble blowing machine, have independent exponential distribution with the same parameter λ. We know that, given a bubble did not burst in 1 second, then the probability that it does not burst in 3 seconds is e^{-2}. Determine λ. If we always generate a new bubble right after the previous one burst, what is the probability that the total life span of 100 bubbles is more than 2 minutes? *Solution:*

(2 points) Let X_i be the time taken for the i^{th} bubble to burst. Then $X_1, X_2, ..., X_{100}$ are i.i.d random variables with $Exp(\lambda)$ distribution.
(1 point) For any X_i, we further know that $P(X_i > 3|X_i > 1) = e^{-2}$.
(2 points) But since, exponential distribution is memoryless,
(2 points) $P(X_i > 2) = e^{-2}$,
(1 point) so, $\lambda = 1$.

(2 points) We know that $E(X_i) = 1$ and $\sigma_{X_i} = 1$, so
(2 points) $E(\sum_{i=1}^{100} X_i) = 100$ and $\sigma_{\sum_{i=1}^{100} X_i} = 10$.
(1 point) Let $X = \frac{\sum_{i=1}^{100} X_i - 100}{10}$
(1 point) Then by the Central Limit Theorem,
(1 point) $X \sim N(0, 1)$.
(1 point) We are required to find $P(\sum_{i=1}^{100} X_i > 120)$
(2 points) Standardizing, this is $P(X > \frac{120-100}{10}) = P(X > 2) = 1 - \Phi(2) = 0.0228$.
Since no Phi table was provided, answers needn't be numeric to get full marks.

5. We are given the $[0, 1] \times [0, 1]$ square and we uniformly, randomly pick a point in it. We also independently and uniformly pick a number c in the interval $[1, 11]$ and define a curve $y = x^c, 0 < x < 1$.

(a) Given that $c = c_0$, what is the probability that the point is above this curve?
(b) What is the probability that the point is above the curve $y = x^c$? (You may use that $ln(6) = 1.7918$).
Solution:

(3 points) Good picture
(1 point) We have to find the area above the curve $y = x^{c_0}$.
(3 points) This is $\int_0^1 1 - x^{c_0} dx = 1 - \frac{1}{c_0+1} = \frac{c_0}{c_0+1}$

(1 points) Let A be the event that the point is above the curve.
(2 points) The density function of c is $f_c(x) = \frac{1}{10}, 1 \leq x \leq 11$
(1 points) Using Law of Total Probability,
(3 points) $P(A) = \int_{-\infty}^{\infty} P(A|c = x) f_c(x) dx$
(3 points) $= \int_1^{11} \frac{x}{x+1} \frac{1}{10} dx$
(3 points) solving which we get $= 1 - \frac{ln(6)}{10} = 0.82082$

6.* Let X and Y be random variables such that their moment generating functions are $M_X(t) = e^{2t+\frac{5}{2}t^2}$ and $M_Y(t) = e^{t+\frac{1}{2}t^2}$. We know that $Cov(X, Y) = -2$. Find the joint probability density function $f_{X,Y}$. What is the probability $\mathbb{P}(2X < Y + 5)$?
Solution:

(1 points) Since MGFs uniquely determine the distribution of the random variables and here the MGF is of the form of the normal distribution,
(4 points) $X \sim N(2, 5)$ and $Y \sim N(1, 1)$.
(1 points) Let $\boldsymbol{X}$ be a bivariate normal distribution vector and we assume that its components are X and Y.
(1+2 points) Then, $\boldsymbol{\mu_X} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}$.
(1+2 points) Then $det(\boldsymbol{\Sigma}) = 1$ and its inverse is $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$.
(0 points) Then, the joint density function of $\boldsymbol{X}$ is,
(3 points) $\frac{1}{2\pi} e^{-\frac{1}{2}((x-2)^2+4(x-2)(y-1)+5(y-5)^2)}$.

(1 points) To find $P(2X < Y + 5)$ we can rewrite it as $P(2X - Y < 5)$.
(1+2 points) Since X, Y are components of a bivariate normal distribution, $2X - Y \sim N(3, 29)$,
(1 points) so, $P(2X - Y < 5) = \Phi(\frac{2}{\sqrt{29}})$.