IE 230 Seat # ________ Name ___ < KEY > ___

Please read these directions.

Closed book and notes. 60 minutes.

Covers through the normal distribution, Section 4.7 of Montgomery and Runger, fourth edition.

Cover page and four pages of exam.
Page 8 of the Concise Notes.

No calculator. No need to simplify beyond probability concepts.
For example, unsimplified factorials, integrals, sums, and algebra receive full credit.

Throughout, f denotes probability mass function or probability density function
and F denotes cumulative distribution function.

A true-or-false statement is true only if it always true; any counter-example makes it false.

For one point, write your name neatly on this cover page and circle your family name.

Score ____________________________

Closed book and notes. 60 minutes.

For statements a—g, choose true or false, or leave blank.
(three points if correct, one point if left blank, zero points if incorrect)

(a) T F ← The continuous-uniform family of distributions is a special case of the discrete-uniform family.

(b) T F ← Scheduled arrivals, such as to physician's office, are naturally modeled with a Poisson process.

(c) T F ← Consider a discrete random variable X with probability mass function f_X. Then $\int_{-\infty}^{\infty} f_X(c)\,dc = 1$.

(d) T F ← Consider a random variable X with Poisson distribution with mean $\mu = 3$ arrivals. Because $\sigma_X^2 = \mu$, the units of σ_X are arrivals$^{1/2}$.

(e) T ← F If Z is a standard-normal random variable, then $\mathrm{P}(Z = 0) = 0$.

(f) T F ← If Z is a standard-normal random variable, then $f_Z(0) = 0$, where f_Z is the probability density function of Z.

(g) T F ← If Z is a standard-normal random variable, then $F_Z(-3.2) = F_Z(3.2)$, where F_Z is the cdf of Z.

2. Suppose that the random variable X has the discrete uniform distribution over the set $\{1, 2, ..., 10\}$ and that the random variable Y has the continuous uniform distribution over the set $[1, 10]$.

For statements a–e, choose true or false.

(a) (3 points) T ← F $\mathrm{E}(X) = \mathrm{E}(Y)$.

(b) (3 points) T F ← $\mathrm{V}(X) = \mathrm{V}(Y)$.

(c) (3 points) T F ← $F_X(5.5) = F_Y(5)$.

(d) (3 points) T F ← $f_X(5.5) = f_Y(5)$.

(e) (3 points) T ← F $\mathrm{P}(X = 5) = f_X(5)$.

3. (from Montgomery and Runger, 4–50) For a battery in a laptop computer under common conditions, the time from being fully charged until needing to be recharged is normally distributed with mean 260 minutes and a standard deviation of 50 minutes.

(a) (8 points) Sketch (well) the corresponding normal pdf. Label and scale both axes.

Sketch the usual bell curve, with center at 260 minutes
and points of inflection at 210 and 310 minutes.
Label the horizonal axis with a dummy variable, such as x.
Label the vertical axis with $f_X(x)$.
Scale the horizonal axis with at least two numbers, such as 260 and 310.
Scale the vertical axis with at least two numbers, such as zero and
the mode height, which is $1/(\sqrt{2\pi}\sigma) \approx (0.4)/(50) = 0.008$

(b) (6 points) In your sketch, show the probability that a randomly selected recharge is less than three hours.

(c) (5 points) State the numerical value of the probability in Part (b). State as much precision as you can determine (given that you don't have access to a normal table).

For Part (b) shade the area under f_X to the left of 180 minutes.
That probability is $\mathrm{P}(X < 180) = \mathrm{P}(Z < (180 - 260)/50) = \mathrm{P}(Z \le -1.6)$
We know that $\mathrm{P}(Z \le -2) \approx 0.05$ and $\mathrm{P}(Z \le -1) \approx 0.16$.
Therefore, $\mathrm{P}(X < 180) \approx 0.1 \leftarrow$

(d) (4 points) Give at least one reason why the time until recharge cannot possibly have a normal distribution.

The range of the normal distribution is $(-\infty, \infty)$.
The range of time to recharge is $(0, \infty)$.

---

4. Consider the binomial pmf $f_X(x) = C_x^n\, p^x\, (1-p)^{n-x}$ for $x = 0, 1, ..., n$ and zero elsewhere.

(a) (5 points) Explain, in words, the origins of p^x. Include any assumptions that are required for your explanation.

The probability of x independent trials all being successful, when p is the probability of success for each trial.

(b) (3 points) Evaluate C_3^8.

$$C_3^8 = \frac{8!}{3!\,5!} = \frac{8 \times 7 \times 6}{3 \times 2} = 56 \leftarrow$$

(c) (5 points) Explain, in words, the phrase "and zero elsewhere".

The pmf f_X is defined for every real number x.
The interesting part of f_X, those values of x that are possible, are defined explicitly.
The uninteresting part of f_X, those values that are impossible, are defined implicitly with the phrase "and zero elsewhere".

---

5. Consider Question 1, which is composed of seven true-false question. Suppose that a clueless student is taking this exam and answers all true-false questions by flipping a coin, with heads yielding "true" and tails yielding "false". Let X denote the number of questions answered correctly.

(a) (6 points) Choose an appropriate distribution for X. Include the family name, parameter values, and probability mass-or-density function f_X.

---

binomial with $n = 7$ and $p = 0.5 \leftarrow$
The pdf is $f_X(x) = C_x^n\, p^x\, (1-p)^{n-x}$ for $x = 0, 1, ..., n$ and zero elsewhere. $\leftarrow$

(b) (3 points) Consider another clueless student who leaves all seven questions blank. Determine this student’s expected number of points?

The mean is 7 points, since the student always gets 7 points.

(c) (3 points) Again consider the student who leaves all seven questions blank. Determine the standard deviation of this student’s number of points.

The std is zero points, since always the student always receives seven points. $\leftarrow$

---

6. (from Montgomery and Runger, 4–11) Suppose that the cumulative distribution function of the random variable X is $F_X(x) = 0.2x$ for $0 \le x \le c$. Assume that values outside the interval $[0, c]$ are not possible.

(a) (5 points) Sketch F_X over the entire real-number line. Label and scale both axes.

Sketch two perpendicular axes.
Label the horizontal axis with a dummy variable, such as x.
Label the vertical axis with $F_X(x)$.
Scale the horizontal axis with at least two numbers, probably zero and c.
Scale the vertical axis with at least two numbers, probably zero and one.
Plot F_X, showing the values for all real numbers x.

(b) (5 points) Determine the value of c.

$F_X(c) = \mathrm{F}_X(c) = 1$ at the upper bound, so
$c = 5 \leftarrow$

(c) (5 points) Write f_X. Be complete.

The density function is the first derivative of the cdf, so
$f_X(x) = 1/5$ for $0 \le x \le 5$ and is zero elsewhere. $\leftarrow$

Discrete Distributions: Summary Table

random variable	distribution name	range	probability mass function	expected value	variance
X	general	$x_1, x_2, \ldots, x_n$	$\mathrm{P}(X = x)$ $= f(x)$ $= f_X(x)$	$\sum_{i=1}^{n} x_i f(x_i)$ $= \mu = \mu_X$ $= \mathrm{E}(X)$	$\sum_{i=1}^{n} (x_i - \mu)^2 f(x_i)$ $= \sigma^2 = \sigma_X^2$ $= \mathrm{V}(X)$ $= \mathrm{E}(X^2) - \mu^2$
X	discrete uniform	$x_1, x_2, \ldots, x_n$	$1/n$	$\sum_{i=1}^{n} x_i / n$	$[\sum_{i=1}^{n} x_i^2 / n] - \mu^2$
X	equal-space uniform	$x = a, a+c, \ldots, b$ where	$1/n$ $n = (b-a+c)/c$	$\dfrac{a+b}{2}$	$\dfrac{c^2(n^2-1)}{12}$
"# successes in 1 Bernoulli trial"	indicator variable	$x = 0, 1$	$p^x (1-p)^{1-x}$	p where	$p(1-p)$ $p = \mathrm{P}(\text{"success"})$
"# successes in n Bernoulli trials"	binomial	$x = 0, 1, \ldots, n$	$C_x^n \, p^x (1-p)^{n-x}$	np where	$np(1-p)$ $p = \mathrm{P}(\text{"success"})$
"# successes in a sample of size n from a population of size N containing K successes"	hyper-geometric (sampling without replacement)	$x =$ $(n-(N-K))^+$, ..., $\min\{K, n\}$ and integer	$C_x^K \, C_{n-x}^{N-K} / C_n^N$	np where	$np(1-p)\dfrac{(N-n)}{(N-1)}$ $p = K/N$
"# Bernoulli trials until 1st success"	geometric	$x = 1, 2, \ldots$	$p(1-p)^{x-1}$	$1/p$ where	$(1-p)/p^2$ $p = \mathrm{P}(\text{"success"})$
"# Bernoulli trials until rth success"	negative binomial	$x = r, r+1, \ldots$	$C_{r-1}^{x-1} \, p^r (1-p)^{x-r}$	r/p where	$r(1-p)/p^2$ $p = \mathrm{P}(\text{"success"})$
"# of counts in time t from a Poisson process with rate λ"	Poisson	$x = 0, 1, \ldots$	$e^{-\mu} \mu^x / x!$	μ where	μ $\mu = \lambda t$

Result. For $x = 1, 2, \ldots$, the geometric cdf is $F_X(x) = 1 - (1-p)^x$.

Result. The geometric distribution is the only discrete memoryless distribution.
That is, $\mathrm{P}(X > x + c \mid X > x) = \mathrm{P}(X > c)$.

Result. The binomial distribution with $p = K/N$ is a good approximation to the hypergeometric distribution when n is small compared to N.

Continuous Distributions: Summary Table

random variable	distribution name	range	cumulative distrib. func.	probability density func.	expected value	variance
X	general	$(-\infty, \infty)$	$P(X \le x)$ $= F(x)$ $= F_X(x)$	$\left.\frac{dF(y)}{dy}\right\vert_{y=x}$ $= f(x)$ $= f_X(x)$	$\int_{-\infty}^{\infty} x f(x) dx$ $= \mu = \mu_X$ $= E(X)$	$\int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$ $= \sigma^2 = \sigma_X^2$ $= V(X)$ $= E(X^2) - \mu^2$
X	continuous uniform	$[a, b]$	$\frac{x-a}{b-a}$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
X	triangular	$[a, b]$	$(x-a)f(x)/2$ if $x \le m$, else $1-(b-x)f(x)/2$	$\frac{2(x-d)}{(b-a)(m-d)}$ $(d = a$ if $x \le m$,	else $d = b)$ $\frac{a+m+b}{3}$	$\frac{(b-a)^2-(m-a)(b-m)}{18}$
sum of random variables	normal (or Gaussian)	$(-\infty, \infty)$	Table III	$\frac{e^{\frac{-1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}}{\sqrt{2\pi}\,\sigma}$	μ	σ^2
time to Poisson count 1	exponential	$[0, \infty)$	$1 - e^{-\lambda x}$	$\lambda e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$
time to Poisson count r	Erlang	$[0, \infty)$	$\sum_{k=r}^{\infty} \frac{e^{-\lambda x}(\lambda x)^k}{k!}$	$\frac{\lambda^r x^{r-1} e^{-\lambda x}}{(r-1)!}$	r/λ	r/λ^2
lifetime	gamma	$[0, \infty)$	numerical	$\frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)}$	r/λ	r/λ^2
lifetime	Weibull	$[0, \infty)$	$1 - e^{-(x/\delta)^\beta}$	$\frac{\beta x^{\beta-1} e^{-(x/\delta)^\beta}}{\delta^\beta}$	$\delta\Gamma(1+\frac{1}{\beta})$	$\delta^2\Gamma(1+\frac{2}{\beta})-\mu^2$

Definition. For any $r > 0$, the *gamma function* is $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x}\, dx$.

Result. $\Gamma(r) = (r-1)\Gamma(r-1)$. In particular, if r is a positive integer, then $\Gamma(r) = (r-1)!$.

Result. The exponential distribution is the only continuous memoryless distribution.
That is, $P(X > x + c \mid X > x) = P(X > c)$.

Definition. A *lifetime* distribution is continuous with range $[0, \infty)$.

Modeling lifetimes. Some useful lifetime distributions are the exponential, Erlang, gamma, and Weibull.