IE 230 Seat # ________ (1 pt.) Name ________________________

Closed book and notes. 60 minutes.

Cover page and four pages of exam.
Pages 8 and 12 of the Concise Notes.

No calculators.

This test covers through Chapter 4 of Montgomery and Runger, third edition.

Score ____________________________

IE 230 — Probability & Statistics in Engineering I (1 pt.) Name ________________________

Closed book and notes. 60 minutes.

1. True or false. (For each: two points if correct, zero points if blank or wrong.)

(a) T ← F The family of normal probability density functions has only one shape.

(b) T ← F The mode of the exponential distribution is zero.

(c) T F ← The continuity correction is used when approximating a continuous distribution with a discrete distribution.

(d) T F ← The distribution of human lifetimes has the memoryless property.

(e) T ← F The rate λ of a Poisson process depends upon the units chosen for time.

(f) T F ← $\Gamma(5) = 120$.

(g) T ← F If F_X denotes the cdf of the random variable X,

$$P(10 < X \leq 14) = F_X(14) - F_X(10).$$

2. (2 pts. each) Write the name of the family of distributions that matches the stated property.

(a) Number of successes in n Bernoulli trials. __ < binomial > __

(b) Number of Bernoulli trials until the rth success. __ < neg. binomial > __

(c) Number of occurrences in a given interval of a Poisson process. __ < Poisson > __

(d) The result of MSexcel's "rand()" function. __ < cont. uniform > __

(e) Discrete distributions with the memoryless property. __ < geometric > __

(f) Time until the rth occurrence in a Poisson process. __ < Erlang > __

(g) When sampling without replacement, the number of successes in n draws from a population of size N that contains K successes. __ < hypergeometric > __

(h) All possible values are equally likely. __ < uniform > __

(i) The "bell curve". __ < normal > __

(j) Generalization of Erlang distributions. __ < gamma > __

3. Suppose that exam scores are normally distributed with a mean of 73 and standard deviation of 9.

   (a) (6 pts.) Sketch the pdf. Label and scale both axes.

   Label the horizontal axis with a dummy variable, such as x.
   Sketch a bell curve.
   Scale the horizontal axis with 73 under the center of the bell curve
   and64 and 82 at the points of inflection.
   Label the vertical axis with $f_X(x)$.
   Scale the vertical axis with zero at the bottom and about $0.4/9 \approx 0.044$
   at the top of the bell curve.

   (b) (4 pts.) Determine the standard-normal z-score that corresponds to an exam score of 60?

   $z = (x - \mu) / \sigma = (60 - 73) / 9 = -13 / 9 \leftarrow$

   (c) (5 pts.) Determine the probability of scoring above 100?

   $\mathrm{P}(X > 100) \approx 0.0015 \leftarrow$

   (Recall: $\mathrm{P}(\mu - 3\sigma < X < \mu + 3\sigma) \approx 0.997$.)

   ---

4. (Montgomery and Runger, 4–80) Assume that the times between arrivals of email notes to your account are exponentially distributed with mean of 23 minutes. Also assume that times are independent of each other.

   (a) (3 pts.) Determine the arrival rate λ. (including units)

   $\lambda = 1 / \mu = 1 / 23$ notes per minute.

   (b) (3 pts.) Determine the standard deviation of the number of arrivals in one hour.

   Let X denote the number of arrivals in one hour, which is 60 minutes.
   Then X is Poisson with mean $t\,\lambda = 60/23$ notes.
   From Page 12, $\mathrm{V}(X) = \mathrm{E}(X) = 60/23$, so $\mathrm{std}(X) = \sqrt{60/23} \leftarrow$

   (c) (3 pts.) Determine the standard deviation of the time until two notes arrive.

   Let Y denote the time until two more notes arrive.
   Then Y is Erlang with parameters $r = 2$ and λ.
   From Page 12, $\mathrm{V}(Y) = r / \lambda^2 = (2)(23)^2$.
   Therefore, $\mathrm{std}(Y) = 23\sqrt{2} \leftarrow$.

   ---

5. (Montgomery and Runger, 4–115) Suppose that X has a Weibull distribution with $\beta = 1$ and $\delta = 1000$.

(a) (2 pts.) What is another name for this distribution?

---

exponential ←

(b) (2 pts.) T ← F This is a lifetime distribution.

(c) (2 pts.) T F ← $P(X > 10 \mid X > 4) = P(X > 14)$.

(d) (4 pts.) Determine the value of $E(X)$.

$E(X) = \delta\,\Gamma(1 + (1/\beta) = (1000)\Gamma(2) = 1000$ ←

---

6. (Montgomery and Runger, 4–100) In a data-communication system, five messages that arrive at a node are bundled into a packet before they are transmitted over the network. Assume that messages arrive according to a Poisson process with rate $\lambda = 30$ messages per minute.

(a) (3 pts.) Determine the mean time until a packet is formed.

Let X denote the time until a packet is formed.
Then X is Erlang with parameters $r = 5$ and λ.
From Page 12, $E(X) = r / \lambda = 5 / 30 = 1 / 6$ minutes ←

(b) (4 pts.) Determine the probability that a packet is formed in less than ten seconds.

The expected number in ten seconds is $\lambda t = (30)(1/6) = 5$ messages.
Let X be as defined in Part (a).
$P(X < 1/6) = F_X(1/6) = \sum_{x=r}^{\infty} e^{-5} 5^x / x!$.
Computationally, better to use $P(X < 1/6) = 1 - \sum_{x=0}^{r-1} e^{-5} 5^x / x!$.

---

7. (Montgomery and Runger, 4–9) The probability density function of the length (in meters) of a random metal rod is $f_X(y) = c$ for $2.3 < y < 2.8$ and zero elsewhere.

(a) (4 pts.) Determine the value of the constant c.

The value of c should make the pdf integrate to one.
Therefore, $c = 1 / (2.8 - 2.3) = 2 \leftarrow$

(b) (6 pts.) Determine the cumulative distribution function, F_X.

In general, for every real number x,

$$F_X(x) = \mathrm{P}(X \le x) = \int_{-\infty}^{x} f_X(y)\, dy .$$

Therefore,

$F_X(x) = 0$ if $x < 2.3$.
$F_X(x) = 1$ if $x > 2.8$.
$F_X(x) = (x - 2.3) / (2.8 - 2.3) = 2(x - 2.3)$ if $2.3 \le x \le 2.8$

(c) (4 pts.) Determine the standard deviation of rod lengths.

X is uniformly distributed between $a = 2.3$ and $b = 2.8$ meters.
From Page 12, $\mathrm{V}(X) = (b - a)^2 / 12$, so

$$\mathrm{std}(X) = (b - a) / \sqrt{12} = (2.8 - 2.3) / \sqrt{12} = 0.5 / \sqrt{12} \leftarrow$$

(d) (6 pts.) Determine the probability that a rod is within 0.1 meter of 2.5 meters long.

$$\mathrm{P}(|X - 2.5| < 0.1) = \mathrm{P}(2.4 < X < 2.6) = F_X(2.6) - F_X(2.4)$$
$$= [(2.6 - 2.3) / (2.8 - 2.3)] - [(2.4 - 2.3) / (2.8 - 2.3)] = 0.4 \leftarrow$$

(Alternatively, notice that the distance between 2.4 and 2.6 is 0.2, which is 40% of the distance between 2.3 and 2.8.)

Discrete Distributions: Summary Table

random variable	distribution name	range	probability mass function	expected value	variance
X	general	$x_1, x_2, \ldots, x_n$	$P(X = x)$ $= f(x)$ $= f_X(x)$	$\sum_{i=1}^{n} x_i f(x_i)$ $= \mu = \mu_X$ $= E(X)$	$\sum_{i=1}^{n} (x_i - \mu)^2 f(x_i)$ $= \sigma^2 = \sigma_X^2$ $= V(X)$ $= E(X^2) - \mu^2$
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	$1/n$ $n = (b-a+c)/c$	$\frac{a+b}{2}$	$\frac{c^2(n^2-1)}{12}$
	indicator variable	$x = 0, 1$	$p^x (1-p)^{1-x}$	p where	$p(1-p)$ $p = P(\text{"success"})$
	binomial	$x = 0, 1, \ldots, n$	$C_x^n p^x (1-p)^{n-x}$	np where	$np(1-p)$ $p = P(\text{"success"})$
	hyper-geometric	$x =$ $(n-(N-K))^+$, $\ldots, \min\{K, n\}$ and integer	$C_x^K C_{n-x}^{N-K} / C_n^N$	np where	$np(1-p)\frac{(N-n)}{(N-1)}$ $p = K/N$
	geometric	$x = 1, 2, \ldots$	$p(1-p)^{x-1}$	$1/p$ where	$(1-p)/p^2$ $p = P(\text{"success"})$
	negative binomial	$x = r, r+1, \ldots$	$C_{r-1}^{x-1} p^r (1-p)^{x-r}$	r/p where	$r(1-p)/p^2$ $p = P(\text{"success"})$
	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, $P(X > x + c \mid X > x) = P(X > c)$.

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	$(-\infty, \infty)$	$P(X \le x)$ $= F(x)$ $= F_X(x)$	$\left.\frac{dF(y)}{dy}\right\|_{y=x}$ $= f(x)$ $= f_X(x)$	$\int_{-\infty}^{\infty} x f(x) dx$ $= \mu = \mu_X$ $= E(X)$	$\int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$ $= \sigma^2 = \sigma_X^2$ $= V(X)$ $= E(X^2) - \mu^2$
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 II	$\frac{e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sqrt{2\pi}\,\sigma}$	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.