

2008 Probability / Mathematical Statistics Comprehensive Exam (100 Points)

1. Suppose $P(A') = .4$, $P(B') = .7$, and $P(A \cap B) = .2$. Find:

- (a) (2pt) $P(A \cup B)$
- (b) (3pt) $P(B|A)$
- (c) (3pt) $P(B|A')$
- (d) (2pt) Are events A and B independent? Justify your answer.

2. X is a continuous random variable with pdf $f(x) = \begin{cases} (x - 1)/4 & \text{if } 2 \leq x \leq 4 \\ 0 & \text{otherwise} \end{cases}$

- (a) (4pt) Find $E(X)$.
- (b) (4pt) Find the CDF $F(x)$. Be sure to define $F(x)$ for all real numbers.

3. (3pt) In a box there are 10 identical balls except for the colors: 4 are white and 6 are red. If I pick a white ball I win a prize. If I pick a red ball, I pay \$2.00. The ball is replaced after being selected. Let X be the number of times I play until (and including) the second time I win. What is the distribution of X and what are the associated parameter values?

4. (3pt) In a manufacturing process that makes light bulbs, there is a 0.01 probability that any particular bulb produced is defective. Assume this probability is constant over time. The bulbs are put into boxes containing 100 bulbs. Let X be the number of defective bulbs in a randomly selected box of bulbs. What is the distribution of X and what are the associated parameter values?

5. (4pt) X and Y are random variables with joint pdf $f(x, y) = \begin{cases} \frac{1}{2}xy & \text{for } 0 < y < x < 2 \\ 0 & \text{otherwise} \end{cases}$

Find $f_X(x)$, the marginal pdf for X .

6. Two discrete random variables X and Y have joint pdf defined by the following table:

		Y	
		1	2
X	1	.10	.15
	2	.30	.45

- (a) (2pt) What is the marginal pdf $f_Y(y)$?
- (b) (2pt) What is the conditional pdf $f(Y|X = 1)$?
- (c) (2pt) What is $E(Y|X = 1)$?
- (d) (2pt) Are X and Y independent? Justify your answer.

7. (6pt) Find the pdf of $Z = \frac{1}{X}$ if X is a random variables with pdf and CDF

$$f_X(x) = \begin{cases} \frac{2}{x^3} & \text{for } 1 < x < \infty \\ 0 & \text{otherwise} \end{cases} \quad F_X(x) = \begin{cases} 1 - \frac{1}{x^2} & \text{for } 1 < x < \infty \\ 0 & \text{otherwise} \end{cases}$$

8. Suppose $X_1 \sim \text{EXP}(3)$, $X_2 \sim \text{GAMMA}(3, 4)$, and X_1 and X_2 are independent.
- (5pt) What is the moment generating function of $Y = X_1 + X_2$?
 - (3pt) What is the distribution of Y ?
9. (8pt) Verify that \bar{X} is a UMVUE of μ for a Poisson(μ) distribution based on a random sample of size n .
10. Let X_1, X_2, \dots, X_n be a random sample from $f(x; \theta) = \theta x^{\theta-1}$ with support $0 < x < 1$ and with $\theta > 0$. Note this is the Beta($\theta, 1$) pdf.
- (3pt) Find the method of moments estimator (MME) of θ .
 - (6pt) Find $\hat{\theta}$, the maximum likelihood estimator (MLE) of θ .
 - (2pt) What is the MLE of $\frac{1}{\theta + 1}$?
11. (8pt) Suppose X and Y are independent EXP(1) random variables. Find the joint pdf of $S = X + Y$ and $T = X$.
12. (5pt) Find a complete sufficient statistic if we have a random sample X_1, X_2, \dots, X_n from a distribution with pdf $f(x; \theta) = \frac{(\ln \theta)\theta^x}{\theta - 1}$ for $0 < x < 1$ and $\theta > 1$.
13. (7pt) Let X_1, X_2, \dots, X_n be a random sample from a Gamma(2, $1/\theta$) pdf. It can be shown that $2\theta X_i \sim \chi^2(4)$. Determine an equal-tailed $100\gamma\%$ interval estimator for θ based on $Y = \sum_{i=1}^n 2\theta X_i$.
14. (5pt) Suppose X_1, \dots, X_{20} are a random sample from an EXP(2) distribution. Approximate the probability that $\bar{X} \geq 2.5$.
15. (6pt) Suppose X and Y are a random sample from an EXP(1) distribution. Find the joint pdf of $S = X + Y$ and $T = X$.

Special Discrete Distributions

Notation and Parameters	Discrete pdf $f(x)$	Mean	Variance	MGF $M_X(t)$
Binomial				
$X \sim \text{BIN}(n, p)$ $0 < p < 1$ $q = 1 - p$	$\binom{n}{x} p^x q^{n-x}$ $x = 0, 1, \dots, n$	np	npq	$(pe^t + q)^n$
Bernoulli				
$X \sim \text{BIN}(1, p)$ $0 < p < 1$ $q = 1 - p$	$p^x q^{1-x}$ $x = 0, 1$	p	pq	$pe^t + q$
Negative Binomial				
$X \sim \text{NB}(r, p)$ $0 < p < 1$ $r = 1, 2, \dots$	$\binom{x-1}{r-1} p^r q^{x-r}$ $x = r, r+1, \dots$	r/p	rq/p^2	$\left(\frac{pe^t}{1-qe^t}\right)^r$
Geometric				
$X \sim \text{GEO}(p)$ $0 < p < 1$ $q = 1 - p$	pq^{x-1} $x = 1, 2, \dots$	$1/p$	q/p^2	$\frac{pe^t}{1-qe^t}$
Hypergeometric				
$X \sim \text{HYP}(n, M, N)$ $n = 1, 2, \dots, N$ $M = 0, 1, \dots, N$	$\frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$ $x = 0, 1, \dots, n$	nM/N	$n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}$	*
Poisson				
$X \sim \text{POI}(\mu)$ $0 < \mu$	$\frac{e^{-\mu} \mu^x}{x!}$ $x = 0, 1, \dots$	μ	μ	$e^{\mu(e^t - 1)}$
Discrete Uniform				
$X \sim \text{DU}(N)$ $N = 1, 2, \dots$	$1/N$ $x = 1, 2, \dots, N$	$\frac{N+1}{2}$	$\frac{N^2-1}{12}$	$\frac{1}{N} \frac{e^t - e^{t(N+1)}}{1 - e^t}$

*Not tractable.

Special Continuous Distributions

Notation and Parameters	Continuous pdf $f(x)$	Mean	Variance	MGF $M_X(t)$
Uniform				
$X \sim \text{UNIF}(a, b)$ $a < b$	$\frac{1}{b-a}$ $a < x < b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt} - e^{at}}{(b-a)t}$
Normal				
$X \sim N(\mu, \sigma^2)$ $0 < \sigma^2$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-[(x-\mu)/\sigma]^2/2}$	μ	σ^2	$e^{\mu t + \sigma^2 t^2/2}$
Gamma				
$X \sim \text{GAM}(\theta, \kappa)$ $0 < \theta$ $0 < \kappa$	$\frac{1}{\theta^\kappa \Gamma(\kappa)} x^{\kappa-1} e^{-x/\theta}$ $0 < x$	$\kappa\theta$	$\kappa\theta^2$	$\left(\frac{1}{1-\theta t}\right)^\kappa$
Exponential				
$X \sim \text{EXP}(\theta)$ $0 < \theta$	$\frac{1}{\theta} e^{-x/\theta}$ $0 < x$	θ	θ^2	$\frac{1}{1-\theta t}$
Two-Parameter Exponential				
$X \sim \text{EXP}(\theta, \eta)$ $0 < \theta$	$\frac{1}{\theta} e^{-(x-\eta)/\theta}$ $\eta < x$	$\eta + \theta$	θ^2	$\frac{e^{\eta t}}{1-\theta t}$
Double Exponential				
$X \sim \text{DE}(\theta, \eta)$ $0 < \theta$	$\frac{1}{2\theta} e^{- x-\eta /\theta}$	η	$2\theta^2$	$\frac{e^{\eta t}}{1-\theta^2 t^2}$
Weibull				
$X \sim \text{WEI}(\theta, \beta)$ $0 < \theta$ $0 < \beta$	$\frac{\beta}{\theta^\beta} x^{\beta-1} e^{-(x/\theta)^\beta}$ $0 < x$	$\theta \Gamma\left(1 + \frac{1}{\beta}\right)$	$\theta^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right]$	*
Extreme Value				
$X \sim \text{EV}(\theta, \eta)$ $0 < \theta$	$\frac{1}{\theta} \exp\left\{ -\left[\frac{x-\eta}{\theta}\right] - \exp\left[\frac{x-\eta}{\theta}\right] \right\}$	$\eta - \gamma\theta$ $\gamma \doteq 0.5772$ (Euler's const.)	$\frac{\pi^2 \theta^2}{6}$	$e^{\eta t} \Gamma(1 + \theta t)$

Special Continuous Distributions

Notation and Parameters	Continuous pdf $f(x)$	Mean	Variance	MGF $M_X(t)$
Cauchy				
$X \sim \text{CAU}(\theta, \eta)$ $0 < \theta$	$\frac{1}{\theta\pi\{1 + [(x - \eta)/\theta]^2\}}$	**	**	**
Pareto				
$X \sim \text{PAR}(\theta, \kappa)$ $0 < \theta$ $0 < \kappa$	$\frac{\kappa}{\theta(1 + x/\theta)^{\kappa+1}}$ $0 < x$	$\frac{\theta}{\kappa - 1}$ $1 < \kappa$	$\frac{\theta^2\kappa}{(\kappa - 2)(\kappa - 1)^2}$ $2 < \kappa$	**
Chi-Square				
$X \sim \chi^2(v)$ $v = 1, 2, \dots$	$\frac{1}{2^{v/2}\Gamma(v/2)} x^{v/2-1} e^{-x/2}$ $0 < x$	v	$2v$	$\left(\frac{1}{1-2t}\right)^{v/2}$
<hr/>				
Notation and Parameters	Continuous pdf $f(x)$	Mean	Variance	MGF $M_X(t)$
Student's t				
$X \sim t(v)$ $v = 1, 2, \dots$	$\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{v\pi}} \left(1 + \frac{x^2}{v}\right)^{-\frac{v+1}{2}}$	0 $1 < v$	$\frac{v}{v-2}$ $2 < v$	**
Snedecor's F				
$X \sim F(v_1, v_2)$ $v_1 = 1, 2, \dots$ $v_2 = 1, 2, \dots$	$\frac{\Gamma\left(\frac{v_1 + v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} x^{\frac{v_1}{2}-1} \times \left(1 + \frac{v_1}{v_2}x\right)^{-\frac{v_1+v_2}{2}}$	$\frac{v_2}{v_2 - 2}$ $2 < v_2$	$\frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)}$ $4 < v_2$	**
Beta				
$X \sim \text{BETA}(a, b)$ $0 < a$ $0 < b$	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$ $0 < x < 1$	$\frac{a}{a+b}$	$\frac{ab}{(a+b+1)(a+b)^2}$	*

*Not tractable.

**Does not exist.

Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Probability / Math Stat Handout to Accompany Exam

- If T is an unbiased estimator of $\tau(\theta)$, then the **Cramer-Rao lower bound (CRLB)** (given a random sample) is

$$\text{Var}(T) \geq \frac{[\tau'(\theta)]^2}{nE\left[\frac{\partial}{\partial\theta} \ln f(X; \theta)\right]^2}$$

assuming (i) the derivatives exist and (ii) they can be passed into the integral (continuous X) or summation (discrete X) when calculating expectations.

- Let statistic $S_i = \mathcal{S}_i(X_1, \dots, X_n)$ and $s_i = \mathcal{S}_i(x_1, \dots, x_n)$ for $i = 1, \dots, k$. Let $\mathcal{S}(x_1, \dots, x_n)$ be the vector of functions whose j^{th} coordinate is $\mathcal{S}_j(x_1, \dots, x_n)$. Then, mathematically, if the conditional pdf of \mathbf{X} given $\mathbf{S} = \mathbf{s}$,

$$f_{\mathbf{X}|\mathbf{s}}(x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n; \theta)}{f_{\mathbf{S}}(\mathbf{s}; \theta)} \quad \text{if } \mathcal{S}(x_1, \dots, x_n) = \mathbf{s} \quad \text{and 0 otherwise}$$

does not depend on θ , then \mathbf{S} is a set of sufficient statistics.

- **Theorem 10.2.1 (Factorization Criterion):** If X_1, \dots, X_n is a random sample having joint pdf $f(x_1, \dots, x_n; \theta)$, and if $\mathbf{S} = (S_1, \dots, S_k)$, then \mathbf{S} is jointly sufficient for θ if and only if

$$f(x_1, \dots, x_n; \theta) = g(\mathbf{s}; \theta) h(x_1, \dots, x_n)$$

where $g(\mathbf{s}; \theta)$ does not depend on observations x_1, \dots, x_n except through \mathbf{s} and $h(x_1, \dots, x_n)$ does not involve θ .

- **Theorem 10.3.1:** If S_1, \dots, S_k are jointly sufficient for θ and if $\hat{\theta}$ is a unique MLE of θ , then $\hat{\theta}$ is a function of $\mathbf{S} = (S_1, \dots, S_k)$.
- A density function is a member of the **regular exponential class (REC)** if it can be expressed in the form

$$f(x; \theta) = c(\theta)h(x) \exp\left[\sum_{j=1}^k q_j(\theta)t_j(x)\right] \quad \text{for } x \in A \text{ (support)}$$

and is 0 otherwise, where $\theta = (\theta_1, \dots, \theta_k)$ is a vector of unknown parameters and the parameter space Ω has the form

$$\Omega = \{\theta \mid a_i < \theta_i < b_i, i = 1, \dots, k\}$$

(Note: $a_i = -\infty$ and $b_i = \infty$ are allowable values.)

and regularity conditions are satisfied.

- **Theorem 10.4.2:** If X_1, \dots, X_n is a random sample from a member of the REC(q_1, \dots, q_k), then the statistics $S_1 = \sum_{i=1}^n t_1(X_i)$, $S_2 = \sum_{i=1}^n t_2(X_i)$, \dots , $S_k = \sum_{i=1}^n t_k(X_i)$ are a minimal set of complete sufficient statistics for $\theta_1, \dots, \theta_k$.