Probability & Statistics for EECS
Fall 2022
Final Exam
December 29, 2022
Time Limit: 180 Minutes

Name (Print): 杨润康

Advisor Name ____________________

This exam contains 12 pages (including this cover page) and 9 problems. Check to see if any pages are missing. Enter all requested information on the top of this page, and put your initials on the top of every page, in case the pages become separated.

Try to answer as many problems as you can. The following rules apply:

- **Mysterious or unsupported answers will NOT receive full credit**. Correct answers, unsupported by calculations, explanation, or algebraic work will receive no credit. Incorrect answers supported by substantially correct calculations and explanations might still receive partial credit.

Do not write in the table to the right.

Problem	Points	Score
1	10	
2	10	
3	10	
4	15	
5	10	
6	10	
7	15	
8	10	
9	10	
Total:	100	

1. (10 points) Basic Concepts.
   (a) (5 points) Please describe the differences and connections between *Probability* and *Statistics*. Then explain why we say conditioning is the soul of our course.
   (b) (5 points) Please describe the pros and cons of *Bayesian statistical inference* and *Classical statistical inference*. Then explain why conjugate priors are important for Bayesian statistical inference.

(a) probability: mathematics likelihood theoretical foundation
study of random events.
statistics: practice of collecting, analyzing, interpreting data
applying probability theory to real-word data.
make predictions about a larger population.
connection: probability provides theoretical basis for statistical methods.
statistics apply these theories to analyze

(b) Bayes: pros: allow for prior beliefs into analysis. update inference as new data available
cons: computationally intensive, choice of prior can significantly influence the result.

Classical: pros: more objective. less computationally intensive
cons: relies on large samples size. relies heavily on point estimate.
don't naturally incorporate prior information or beliefs into the analysis.

2. (10 points) Let X and Y be two continuous random variables with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} x + cy^2 & \text{if } 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

   (a) (5 points) Find the value of constant c.

   (b) (5 points) Find the joint probability $P(0 \le X \le 1/2,\ 0 \le Y \le 1/2)$.

3. (10 points) Let X and Y be two continuous random variables with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 6xy & \text{if } 0 \le x \le 1, 0 \le y \le \sqrt{x}, \\ 0 & \text{otherwise.} \end{cases}$$

(a) (5 points) Find the marginal distributions of X and Y. Are X and Y independent?

(b) (5 points) Find $E[X|Y=y]$ and $\text{Var}[X|Y=y]$ for $0 \le y \le 1$.

(a) The supports of X and Y are both $[0,1]$. In this way, we have

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \\ &= \int_0^{\sqrt{x}} 6xydy \\ &= 3xy^2\Big|_{y=0}^{y=\sqrt{x}} \\ &= 3x^2, \end{aligned}$$

and

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx \\ &= \int_{y^2}^{1} 6xydx \\ &= 3yx^2\Big|_{x=y^2}^{x=1} \\ &= 3y - 3y^5. \end{aligned}$$

Therefore,

$$f_X(x) = \begin{cases} 3x^2 & \text{if } 0 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

$$f_Y(y) = \begin{cases} 3y - 3y^5 & \text{if } 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Since $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$, X and Y are not independent.

(b) Since

$$E[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx,$$

to calculate $E[X|Y=y]$, we need to first calculate $f_{X|Y}(x|y)$.

If $y^2 \le x \le 1$,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{2x}{1-y^4}.$$

In this way,

$$f_{X|Y}(x|y) = \begin{cases} \frac{2x}{1-y^4} & \text{if } y^2 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$\begin{aligned} E[X|Y=y] &= \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx \\ &= \int_{y^2}^{1} x\frac{2x}{1-y^4}dx \\ &= \frac{2}{3(1-y^4)}x^3\Big|_{x=y^2}^{x=1} \\ &= \frac{2(1-y^6)}{3(1-y^4)} \\ &= \frac{2}{3}\cdot\frac{1+y^2+y^4}{1+y^2} \end{aligned}$$

Since

$$\text{Var}[X|Y=y] = E[X^2|Y=y] - (E[X|Y=y])^2,$$

to calculate $\text{Var}[X|Y=y]$, we need to first calculate $E[X^2|Y=y]$.

Since

$$\begin{aligned} E[X^2|Y=y] &= \int_{-\infty}^{\infty} x^2 f_{X|Y}(x|y)dx \\ &= \int_{y^2}^{1} x^2\frac{2x}{1-y^4}dx \\ &= \frac{1}{2(1-y^4)}x^4\Big|_{x=y^2}^{x=1} \\ &= \frac{1-y^8}{2(1-y^4)} \\ &= \frac{1+y^4}{2}, \end{aligned}$$

we have,

$$\begin{aligned} \text{Var}[X|Y=y] &= E[X^2|Y=y] - (E[X|Y=y])^2 \\ &= \frac{1+y^4}{2} - \left(\frac{2(1-y^6)}{3(1-y^4)}\right)^2 \\ &= \frac{1+y^4}{2} - \frac{4}{9}\cdot\frac{(1+y^2+y^4)^2}{(1+y^2)^2} \end{aligned}$$

(c) According to the result in question(b), we have

$$E[X|Y] = \frac{2}{3}\cdot\frac{1+Y^2+Y^4}{1+Y^2},$$

$$\text{Var}[X|Y] = \frac{1+Y^4}{2} - \frac{4}{9}\cdot\frac{(1+Y^2+Y^4)^2}{(1+Y^2)^2}.$$

4. (15 points) Given a coin with the probability p of landing heads, where p is *unknown* and we need to estimate its value through data. In our data collection model, we have n independent tosses, result of each toss is either Head or Tail. Let X denote the number of heads in the total n tosses. Now we conduct experiments to collect data and find out that $X = k$. Then we need to find $\widehat{p}$, the estimation of p.

   (a) (5 points) Assume p is an unknown constant. Find $\widehat{p}$ through the MLE (Maximum Likelihood Estimation) rule.

   (b) (5 points) Assume p is a random variable with a prior distribution Beta(a, b), where a and b are known constants. Find $\widehat{p}$ through the MAP (Maximum a Posterior Probability) rule.

   (c) (5 points) Assume p is a random variable with a prior distribution Beta(a, b), where a and b are known constants. Find $\widehat{p}$ through the MMSE (Minimum Mean Square Estimate) rule.

(a) $P_{X_i}(x_i; p) = p^{x_i}(1-p)^{1-x_i} = \begin{cases} p & x_i = 1 \\ 1-p & x_i = 0 \end{cases}$

Likelihood function $P_X(x; p) = \prod_{i=1}^{n} P_{X_i}(x_i; p) = p^k (1-p)^{n-k}$

$g(p) = \log P_X(x; p) = k \log p + (n-k)\log(1-p)$

$g'(p) = \frac{k}{p} - \frac{n-k}{1-p}$

$g''(p) = -\frac{k}{p^2} - \frac{n-k}{(1-p)^2} \leq 0$

let $g'(p) = 0 \Rightarrow p = \frac{k}{n}$ since $g''(p) \leq 0 \Rightarrow \hat{p}_{MLE} = \frac{k}{n}$

(b) $f_{p|X=k} \propto p^{a+k-1}(1-p)^{b+n-k-1}$ $p \in (0,1)$

$g(p) = \log(f_{p|X=k}) = (a+k-1)\log p + (b+n-k-1)\log(1-p)$

$g'(p) = \frac{a+k-1}{p} - \frac{b+n-k-1}{1-p}$

$g''(p) = -\frac{a+k+1}{p^2} - \frac{b+n-k-1}{(1-p)^2} < 0$

$g'(p) = 0 \Rightarrow \hat{p}_{MAP} = \frac{a+k-1}{a+b+n-2}$

(c) $p \mid X=k \sim Beta(a+k, b+n-k)$

$\hat{p}_{MMSE} = E(p \mid X=k) = \frac{a+k}{a+b+n}$

5. (10 points) We know that the MMSE of Y given X is given by $g(X) = E[Y|X]$. We also know that the LLSE (Linear Least Square Estimate) of Y given X, denoted by $L[Y|X]$, is shown as follows:

$$L[Y|X] = E(Y) + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}\big(X - E(X)\big).$$

Now we wish to estimate the probability of landing heads, denoted by θ, of a biased coin. We model θ as the value of a random variable Θ with a known prior with PDF $f_\Theta \sim \mathrm{Unif}(0,1)$. We consider n independent tosses and let X be the number of heads observed.

(a) (5 points) Show that $E\big[(\Theta - E[\Theta|X])h(X)\big] = 0$ for any real function $h(\cdot)$.

(b) (5 points) Find the MMSE $E[\Theta|X]$ and the LLSE $L[\Theta|X]$.
(Eve's law: $\mathrm{Var}(Y) = E\big[\mathrm{Var}(Y|X)\big] + \mathrm{Var}\big[E(Y|X)\big]$.)

(a) $E[(\theta - E[\Theta|X])h(x)]$

$= E[\theta\, h(x) - E[\Theta|X]\, h(x)]$

$= E[\theta\, h(x)] - E[E[\theta|X]\, h(x)]$

$= E[\theta\, h(x)] - E[\theta\, h(x)] = 0$

(b) $\theta|X \sim \mathrm{Beta}(X+1, n-X+1)$

MMSE: $E[\theta|X] = \frac{X+1}{n+2}$

6. (10 points) Laplace's law of succession says that if $X_1, X_2, \ldots, X_{n+1}$ are conditionally independent Bern(p) r.v.s given p, but p is given a Unif$(0,1)$ prior to reflect ignorance about its value, then

$$P\left(X_{n+1}=1 \mid X_1+\cdots+X_n=k\right)=\frac{k+1}{n+2}.$$

As an example, Laplace discussed the problem of predicting whether the sun will rise tomorrow, given that the sun did rise every time for all n days of recorded history; the above formula then gives $(n+1)/(n+2)$ as the probability of the sun rising tomorrow (of course, assuming independent trials with p unchanging over time may be a very unreasonable model for the sunrise problem).

(a) (5 points) Find the posterior distribution of p given $X_1 = x_1, \ldots, X_n = x_n$, and show that it only depends on the sum of the $x_j, j \in \{1, \ldots, n\}$.

(b) (5 points) Prove Laplace's law of succession, using a form of LOTP to find

$$P\left(X_{n+1}=1 \mid X_1+\cdots+X_n=k\right)$$

by conditioning on p. after observing $X_1=x_1$ is Beta $(1+x_1, 1+(1-x_1))$

(a) $p \mid X_1=x_1, \ldots X_n=x_n \sim \text{Beta}\left(1+\sum_{i=1}^{n} x_i,\ 1+n-\sum_{i=1}^{n} x_i\right)$

(b) LOTP. $S_n = X_1+X_2+\cdots+X_n$

$$P(X_{n+1}=1 \mid S_n=k) = \int_0^1 P(X_{n+1}=1 \mid p, S_n=k)\, f(p \mid S_n=k)\, dp$$

$$= \int_0^1 p \cdot f(p \mid S_n=k)\, dp$$

$$= \int_0^1 p \cdot \frac{\Gamma(k+1+n-k+1)}{\Gamma(k+1)\Gamma(n-k+1)} \cdot p^{k+1-1}(1-p)^{n-k+1-1}\, dp$$

$$= \frac{\Gamma(n+2)}{\Gamma(k+1)\Gamma(n-k+1)} \underbrace{\int_0^1 p^{k+2-1}(1-p)^{n-k+1-1}\, dp}_{=\beta(k+2,\, n-k+1) = \frac{\Gamma(k+2)\Gamma(n-k+1)}{\Gamma(n+3)}}$$

$$= \frac{k+1}{n+2}$$

(c) $P(X_{n+1}=1 \mid X_1+\cdots+X_n=k) = E(p \mid X_1+\cdots+X_n=k) = \frac{k+1}{n+2}$

7. (15 points) A handy rule of thumb in statistics and life is as follows: *Conditioning often makes things better.* This problem explores how the above rule of thumb applies to estimating unknown parameters.

Let θ be an unknown parameter that we wish to estimate based on data $X_1, \dots, X_n$ (these are r.v.s before being observed, and then after the experiment they "crystallize" into data).

In this problem, θ is viewed as an unknown constant, and is *not* treated as an r.v. as in the Bayesian statistical inference. Let T_1 be an estimator for θ (this means that T_1 is a function of $X_1, \dots, X_n$ which is being used to estimate θ).

A strategy for improving T_1 (in some problems) is as follows. Suppose that we have an r.v. R such that $T_2 = E(T_1|R)$ is a function of $X_1, \dots, X_n$ (in general, $E(T_1|R)$ might involve unknowns such as θ but then it couldn't be used as an estimator). Also suppose that $P(T_1 = T_2) < 1$, and that $E(T_1^2)$ is finite.

(a) (5 points) Use Jensen's inequality and Adam's Law to show that T_2 is better than T_1 in the sense that the mean squared error is less, *i.e.*,

$$E\big[(T_2 - \theta)^2\big] < E\big[(T_1 - \theta)^2\big].$$

(b) (5 points) The bias of an estimator T for θ is defined to be $b(T) = E(T) - \theta$. An important identity in statistics, a form of the *bias-variance* trade-off, is that the mean squared error equals the variance plus the squared bias:

$$E\big[(T - \theta)^2\big] = \text{Var}(T) + \big(b(T)\big)^2.$$

Use this identity and Eve's law to give an alternative proof of the result from (a).

(c) (5 points) Now suppose that $X_1, \dots, X_n$ are *i.i.d.* with mean θ, and consider the special case $T_1 = X_1, R = \sum_{j=1}^{n} X_j$. Find T_2 in a simplified form, and check that it has a lower mean squared error than T_1 for $n \geq 2$. Also, explain what happens to T_1 and T_2 as $n \to \infty$.

8. (10 points) Let X and Y be two independent random variables satisfying first success distribution FS(p).

   (a) (5 points) Define $Z_1 = X - Y$. Find the PMF of Z_1 and $E(Z_1)$.

   (b) (5 points) Define $Z_2 = \frac{X}{Y}$. Find the PMF of Z_2 and $E(Z_2)$.

$$P(Z_1 = z) = \sum_{k=1}^{\infty} P(X=k) \cdot P(Y = k-z)$$
$$= \sum_{k=1}^{\infty} p(1-p)^{k-1} \cdot p(1-p)^{k-z-1}$$
$$E(Z_1) = E(X) - E(Y) = 0$$

(a) $FS(p)$: $P(X=k) = p(1-p)^{k-1}$ $k = 1, 2, 3 \ldots$

$$P(X=i, Y=j) = P(X=i) \times P(Y=j) = p(1-p)^{i-1} \times p(1-p)^{j-1}$$

$$P(Z_1 = z) = \sum_{i-j=z} P(X=i, Y=j)$$
$$= \sum_{i=1}^{\infty} P(X=i+z) P(Y=i) \qquad E(Z_1) = E(X) - E(Y) = 0$$
$$= \sum_{i=1}^{\infty} p(1-p)^{i+z-1} \times p(1-p)^{i-1}$$

(b) $F(z) = P(Z \leq z) = P\left(\frac{X}{Y} \leq z\right) = P(X \leq zY)$

$$P(X \leq zY) = \sum_{1}^{\infty} pq^{k-1} \qquad \sum_{0}^{zY}$$

9. (10 points) A scientist makes two measurements X, Y, considered to be *i.i.d.* random variables.

   (a) (5 points) If $X, Y \sim \mathcal{N}(0,1)$. Find the correlation between the larger and smaller of the values, *i.e.*, $\mathrm{Corr}\big(\max(X,Y), \min(X,Y)\big)$.

   (b) (5 points) If $X, Y \sim \mathrm{Unif}(0,1)$. Find the correlation between the larger and smaller of the values, *i.e.*, $\mathrm{Corr}\big(\max(X,Y), \min(X,Y)\big)$.

$\max(x,y) + \min(X,Y) = x+y$

$\max(x,y) - \min(X,Y) = |x-y|$

Define $M = \max(X,Y)$ $L = \min(X,Y)$

(a)

$E(M) + E(L) = E(M+L) = E(X+Y) = E(X) + E(Y) = 0$

$E(M) - E(L) = E(M-L) = E(|X-Y|) = \frac{2}{\sqrt{\pi}}$

$\Rightarrow \begin{cases} E(M) = \frac{1}{\sqrt{\pi}} \\ E(L) = -\frac{1}{\sqrt{\pi}} \end{cases}$

$\mathrm{Cov}(M,L) = E(ML) - E(M)E(L) = E(XY) + \frac{1}{\pi} = E(X)E(Y) + \frac{1}{\pi} = \frac{1}{\pi}$

$\mathrm{Var}(M) = \mathrm{Var}(L)$

$\mathrm{Var}(M+L) = \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = 2$

$\mathrm{Var}(M+L) = \mathrm{Var}(M) + \mathrm{Var}(L) + 2\mathrm{Cov}(M,L) = 2\mathrm{Var}(M) + \frac{2}{\pi}$

$\Rightarrow \mathrm{Var}(M) = \mathrm{Var}(L) = 1 - \frac{1}{\pi}$

$\mathrm{Corr}(M,L) = \frac{\mathrm{Cov}(M,L)}{\mathrm{Var}(M)} = \frac{\frac{1}{\pi}}{1-\frac{1}{\pi}}$

(b) $E(M) + E(L) = E(M+L) = E(X+Y) = 1$

$E(M) - E(L) = E(M-L) = E(|X-Y|) = \frac{1}{3}$

$\Rightarrow \begin{cases} E(M) = \frac{2}{3} \\ E(L) = \frac{1}{3} \end{cases}$

$\mathrm{Cov}(M,L) = E(ML) - E(M)E(L) = E(XY) - \frac{2}{9} = \frac{1}{2} \times \frac{1}{2} - \frac{2}{9} = \frac{1}{36}$

$\mathrm{Var}(M+L) = \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}$

$\mathrm{Var}(M+L) = \mathrm{Var}(M) + \mathrm{Var}(L) + 2\mathrm{Cov}(M,L) = 2\mathrm{Var}(M) + \frac{1}{18} = \frac{1}{6} \Rightarrow \mathrm{Var}(M) = \frac{1}{18}$

$\mathrm{Corr}(M,L) = \frac{\mathrm{Cov}(M,L)}{\mathrm{Var}(M)} = \frac{1}{2}$

	Y **discrete**	Y **continuous**
X **discrete**	$P(Y=y\|X=x)=\frac{P(X=x\|Y=y)P(Y=y)}{P(X=x)}$	$f_Y(y\|X=x)=\frac{P(X=x\|Y=y)f_Y(y)}{P(X=x)}$
X **continuous**	$P(Y=y\|X=x)=\frac{f_X(x\|Y=y)P(Y=y)}{f_X(x)}$	$f_{Y\|X}(y\|x)=\frac{f_{X\|Y}(x\|y)f_Y(y)}{f_X(x)}$

	Y **discrete**	Y **continuous**
X **discrete**	$P(X=x)=\sum_y P(X=x\|Y=y)P(Y=y)$	$P(X=x)=\int_{-\infty}^{\infty} P(X=x\|Y=y)f_Y(y)dy$
X **continuous**	$f_X(x)=\sum_y f_X(x\|Y=y)P(Y=y)$	$f_X(x)=\int_{-\infty}^{\infty} f_{X\|Y}(x\|y)f_Y(y)dy$

Figure 1: Bayes' Rule & LOTP.

Table of distributions

Name	Param.	PMF or PDF	Mean	Variance
Bernoulli	p	$P(X=1)=p, P(X=0)=q$	p	pq
Binomial	n,p	$\binom{n}{k}p^k q^{n-k}$, for $k \in \{0,1,\dots,n\}$	np	npq
FS	p	pq^{k-1}, for $k \in \{1,2,\dots\}$	$1/p$	q/p^2
Geom	p	pq^k, for $k \in \{0,1,2,\dots\}$	q/p	q/p^2
NBinom	r,p	$\binom{r+n-1}{r-1}p^r q^n, n \in \{0,1,2,\dots\}$	rq/p	rq/p^2
HGeom	w,b,n	$\frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$, for $k \in \{0,1,\dots,n\}$	$\mu=\frac{nw}{w+b}$	$(\frac{w+b-n}{w+b-1})n\frac{\mu}{n}(1-\frac{\mu}{n})$
Poisson	λ	$\frac{e^{-\lambda}\lambda^k}{k!}$, for $k \in \{0,1,2,\dots\}$	λ	λ
Uniform	$a<b$	$\frac{1}{b-a}$, for $x \in (a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	μ,σ^2	$\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}$	μ	σ^2
Log-Normal	μ,σ^2	$\frac{1}{x\sigma\sqrt{2\pi}}e^{-(\log x-\mu)^2/(2\sigma^2)}, x>0$	$\theta=e^{\mu+\sigma^2/2}$	$\theta^2(e^{\sigma^2}-1)$
Expo	λ	$\lambda e^{-\lambda x}$, for $x>0$	$1/\lambda$	$1/\lambda^2$
Gamma	a,λ	$\Gamma(a)^{-1}(\lambda x)^a e^{-\lambda x}x^{-1}$, for $x>0$	a/λ	a/λ^2
Beta	a,b	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$, for $0<x<1$	$\mu=\frac{a}{a+b}$	$\frac{\mu(1-\mu)}{a+b+1}$
Chi-Square	n	$\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2}$, for $x>0$	n	$2n$
Student-t	n	$\frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2}$	0 if $n>1$	$\frac{n}{n-2}$ if $n>2$

Figure 2: Table of Random Variables.

Table of distributions

Name	Param.	PMF or PDF	Mean	Variance
Bernoulli	p	$P(X=1)=p, P(X=0)=q$	p	pq
Binomial	n, p	$\binom{n}{k}p^k q^{n-k}$, for $k \in \{0,1,\dots,n\}$	np	npq
FS	p	pq^{k-1}, for $k \in \{1,2,\dots\}$	$1/p$	q/p^2
Geom	p	pq^k, for $k \in \{0,1,2,\dots\}$	q/p	q/p^2
NBinom	r, p	$\binom{r+n-1}{r-1}p^r q^n, n \in \{0,1,2,\dots\}$	rq/p	rq/p^2
HGeom	w, b, n	$\frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$, for $k \in \{0,1,\dots,n\}$	$\mu = \frac{nw}{w+b}$	$(\frac{w+b-n}{w+b-1})n\frac{\mu}{n}(1-\frac{\mu}{n})$
Poisson	λ	$\frac{e^{-\lambda}\lambda^k}{k!}$, for $k \in \{0,1,2,\dots\}$	λ	λ
Uniform	$a < b$	$\frac{1}{b-a}$, for $x \in (a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	μ, σ^2	$\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}$	μ	σ^2
Log-Normal	μ, σ^2	$\frac{1}{x\sigma\sqrt{2\pi}}e^{-(\log x-\mu)^2/(2\sigma^2)}, x > 0$	$\theta = e^{\mu+\sigma^2/2}$	$\theta^2(e^{\sigma^2}-1)$
Expo	λ	$\lambda e^{-\lambda x}$, for $x > 0$	$1/\lambda$	$1/\lambda^2$
Gamma	a, λ	$\Gamma(a)^{-1}(\lambda x)^a e^{-\lambda x}x^{-1}$, for $x > 0$	a/λ	a/λ^2
Beta	a, b	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$, for $0 < x < 1$	$\mu = \frac{a}{a+b}$	$\frac{\mu(1-\mu)}{a+b+1}$
Chi-Square	n	$\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2}$, for $x > 0$	n	$2n$
Student-t	n	$\frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2}$	0 if $n > 1$	$\frac{n}{n-2}$ if $n > 2$

Figure 2: Table of Random Variables.