Stats 200 Practice Final

We will not grade the practice final. The instructions below should give you a sense what you can expect during the final.

Instructions:

- Before you start working on the final, please download this pdf so that in case you close the pdf you still have access to it. During the exam you can find this pdf by clicking on "dashboard" and then "final".
- You can submit as many times as you want. We will only grade your last submission. After submitting your solutions, you can find this pdf on gradescope by clicking on "dashboard", then "final" and then "resubmit".
- The timer starts when you open the assignment. You have 180 minutes to work on the final and 15 minutes to scan and upload your solutions. Thus, the timer on gradescope shows a total of 180+15=195 minutes to complete the assignment. We expect you to stop working on the assignment after 180 minutes and start scanning and uploading.
- We strongly prefer you to submit via gradescope. However, if you can't submit on gradescope for some reason, you can send the submission to rdominik@stanford.edu (preferably as one pdf).
- If you have any questions or experience technical difficulties, please join us on Zoom: `https://stanford.zoom.us/j/93165895821?pwd=VVloUGxyRi9zNzBYaFR4TE5qUk5PZz09`, Password: 610244
- Please write neatly: we can't grade what we can't read. To obtain full credit, you must justify your answers. We are primarily interested in your understanding of concepts, so show us what you know.
- No one is expected to answer all of the questions; but everyone is encouraged to try them.
- This final is open book. You may refer to your textbook, your class notes, the scribe notes, or the exercise sheets. You are not allowed to refer to the internet. You may use a basic calculator (CAS calculators are not allowed).
- At the end of this document, you will find a common distributions table of reference.

1 Problem 1

Consider a random variable X with density function $f_{\mu,b}(x) = c\exp\left(-\frac{|x-\mu|}{b}\right)$ for some $b > 0$ and $\mu \in \mathbb{R}$.

(a) Find the value of the constant c.

(b) Derive the mean of X. *Hint: you may want to use that the density of x is symmetric about μ.*

(c) Suppose that $Y \sim \text{Exponential}(\lambda)$ and $Z \sim \text{Unif}\{-1, 1\}$ (i.e., $Z = 1$ with probability $\frac{1}{2}$ and $Z = -1$ with probability $\frac{1}{2}$). Show that ZY is a random variable with density function $f_{0,1/\lambda}$.

Hint: calculate the CDF and then differentiate. For the CDF, we need to consider two cases: $t > 0$ and $t \leq 0$.

Solution.

(a) To make f be a valid density, we need to have that $\int_{-\infty}^{\infty} f_{\mu,b}(x)dx = 1$. Now,

$$\begin{aligned}
\int_{-\infty}^{\infty} \exp\left(-\frac{|x-\mu|}{b}\right)dx &= \int_{-\infty}^{\mu} \exp\left(-\frac{\mu-x}{b}\right)dx + \int_{\mu}^{\infty} \exp\left(-\frac{x-\mu}{b}\right)dx \\
&= b\exp\left(-\frac{\mu-x}{b}\right)\Big|_{-\infty}^{\mu} + (-b)\exp\left(-\frac{x-\mu}{b}\right)\Big|_{\mu}^{\infty} \\
&= b + (-b)\cdot(-1) \\
&= 2b.
\end{aligned}$$

Hence $c = \frac{1}{2b}$.

(b) There are (at least) two ways to prove this. The first option is to directly compute the integral via partial integration, which is a bit more involved. The second option is to use symmetry of the density, which involves less calculations. By symmetry, for $\mu = 0$,

$$\int_{-\infty}^{\infty} xf_{0,b}(x)dx = \int_{-\infty}^{\infty} \frac{x}{2b}e^{-\frac{|x|}{b}}dx = \int_{0}^{\infty} \frac{x}{2b}e^{-\frac{|x|}{b}}dx + \int_{-\infty}^{0} \frac{x}{2b}e^{-\frac{|x|}{b}}dx = 0.$$

In the last step we used that

$$\int_{-\infty}^{0} \frac{x}{2b}e^{-\frac{|x|}{b}}dx = -\int_{0}^{\infty} \frac{x}{2b}e^{-\frac{|x|}{b}}dx.$$

Thus,

$$\int_{-\infty}^{\infty} xf_{\mu,b}(x)dx = \int_{-\infty}^{\infty} xf_{0,b}(x-\mu)dx = \int_{-\infty}^{\infty} (x+\mu)f_{0,b}(x)dx = 0+\mu\int_{-\infty}^{\infty} f_{0,b}(x)dx = \mu.$$

(c) Let F the CDF of ZY. We have that for $t \geq 0$,

$$\begin{aligned}F(t) &= \mathbb{P}(ZY \leq t) \\ &= \mathbb{P}(ZY \leq t | Z = 1)\mathbb{P}(Z = 1) + \mathbb{P}(ZY \leq t | Z = -1)\mathbb{P}(Z = -1) \\ &= \frac{1}{2}\mathbb{P}(Y \leq t) + \frac{1}{2}\mathbb{P}(Y \geq -t) \\ &= \frac{1}{2}F_Y(t) + \frac{1}{2}\end{aligned}$$

and for $t < 0$,

$$\begin{aligned}F(t) &= \mathbb{P}(ZY \leq t) \\ &= \mathbb{P}(ZY \leq t | Z = 1)\mathbb{P}(Z = 1) + \mathbb{P}(ZY \leq t | Z = -1)\mathbb{P}(Z = -1) \\ &= \frac{1}{2}\mathbb{P}(Y \leq t) + \frac{1}{2}\mathbb{P}(Y \geq -t) \\ &= \frac{1}{2}(1 - F_Y(-t))\end{aligned}$$

So the density of ZY is

$$f(t) = F'(t) = \begin{cases} \frac{1}{2}f_Y(t), & t \geq 0 \\ \frac{1}{2}f_Y(-t), & t < 0 \end{cases}$$

Hence, $f(t) = \frac{1}{2}f_Y(|t|) = \frac{\lambda}{2}e^{-\lambda|t|} = f_{0,1/\lambda}(t)$.

2 Problem 2

Kiki is at the carnival. She visits the Wheel of Fortune booth, where she can spin a giant wheel that's divided into several sections; if the wheel lands on the section that's marked with a gold star, then she wins a prize. Kiki wants to know the probability of winning a prize, call it θ. (It's hard to know θ just from looking at the wheel itself: perhaps the wheel is rigged.) She decides to spin the wheel over and over until she wins a prize for the first time, and she counts how many spins it takes. In other words, she observes $X \sim \text{Geom}(\theta)$, where $E(X) = 1/\theta$.

(a) What is the MLE for θ?

(b) What is the Fisher information?

(c) Suppose Kiki has prior beliefs about θ, which she expresses as $\theta \sim \text{Beta}(a, b)$, where $a, b > 0$ are known constants. What is the posterior distribution of θ given $X = x$, and what is the posterior mean $E(\theta | X = x)$?

Solution.

(a) The log-likelihood is $(x-1)\log(1-\theta)+\log(\theta)$, the derivative is $(x-1)/(\theta-1)+1/\theta$, setting it equal to zero gives us $\hat{\theta}=1/X$.

(b) The second derivative of the log likelihood is $-(x-1)/(1-\theta)^2-1/\theta^2$, so the Fisher information is

$$E(-l''(\theta)) = \frac{\frac{1}{\theta}-1}{(1-\theta)^2}+\frac{1}{\theta^2} = \frac{1}{\theta(1-\theta)}+\frac{1}{\theta^2} = \frac{1}{\theta^2(1-\theta)}.$$

(c) The posterior is proportional to

$$\pi(\theta|x) \propto \theta^{a-1}(1-\theta)^{b-1}\theta(1-\theta)^{x-1} = \theta^{a}(1-\theta)^{b+x-2},$$

which is proportional to the $\text{Beta}(a+1, b+x-1)$ density, so the posterior distribution is $\theta|X=x \sim \text{Beta}(a+1, b+x-1)$. The mean of a $\text{Beta}(a, b)$ is $a/(a+b)$, so the posterior mean is $(a+1)/(a+1+b+x-1) = (a+1)/(a+b+x)$.

3 Problem 3: Hypothesis testing

Suppose we observe one sample X from the following density:

$$\frac{1}{2}p(x-\theta)+\frac{1}{2}p(x+\theta)$$

where $p(x)$ is the standard normal density.

a) What is the distribution of X when $\theta=0$.

b) Fix $\theta_1>0$. Derive the most powerful test for testing $H_0:\theta=0$ vs $H_0:\theta=\theta_1$. *Hint: you may want to use that for $x>0$, the function $f(x)=e^x+e^{-x}$ is increasing in x.*

c) Is it uniformly most powerful for testing $H_0:\theta=0$ vs $H_0:\theta>0$?

Solution.

a) When $\theta=0$, the density is $2\times\frac{1}{2}p(x)=p(x)$. So $X\sim\mathcal{N}(0,1)$.

b) The likelihood ratio is:

$$r(X) = \frac{\frac{1}{2}\exp(-\frac{1}{2}(X-\theta)^2) + \frac{1}{2}\exp(-\frac{1}{2}(X+\theta)^2)}{\exp(-\frac{1}{2}X^2)}$$
$$= \frac{1}{2}\exp\left(-\frac{1}{2}\theta^2\right)(\exp(+X\theta) + \exp(-X\theta))$$
$$= \frac{1}{2}\exp\left(-\frac{1}{2}\theta^2\right)(\exp(+|X|\theta) + \exp(-|X|\theta))$$

$r(X)$ is increasing in $T(X) = |X|$ for $\theta \neq 0$. By Neyman-Pearson, the most powerful test rejects for $T(X) > C$, ie $|X| > C$. Under the null, $X \sim \mathcal{N}(0,1)$ (cf. a)). Therefore, we reject H_0 if $X > z_{\alpha/2}$, where $z_{\alpha/2}$ is the $\alpha/2$-upper quantile of the standard normal distribution (which is the same quantity as the $1-\alpha/2$ quantile of that same distribution).

c) The test derived in b) does not depend on the fixed alternative $\theta_1 > 0$. Therefore that test is uniformly most powerful for testing $H_0 : \theta = 0$ vs $H_0 : \theta > 0$.

4 Problem 4: Testing in tables

Suppose that $n = 1{,}000$ professional athletes are randomly surveyed and asked how many of their feet are infected with athletes foot. Let X_i denote the number of infected feet that the ith subject reported. Their responses, as well as the sex of the subjects, are summarized in the below contingency table[1]:

	$X_i = 0$	$X_i = 1$	$X_i = 2$
Female	357	29	19
Male	513	52	30

(a) Conduct a Pearson chi-squared test of independence to determine whether sex and and the number of feet infected with athletes foot are independent. You must report the chi-squared test statistic, the rejection threshold of the test at level $\alpha = 0.05$, and whether or not the test rejects the null at level $\alpha = 0.05$.

(b) Ignoring sex, the 1,000 X_i samples are summarized in the following 3×1 contingency table

	$X_i = 0$	$X_i = 1$	$X_i = 2$
Count	870	81	49

[1] This is data is entirely synthetic and for educational purposes only. No conclusions about athletes foot or male and female athletes should be drawn from it.

Suppose $X_1, \dots, X_{1000}$ are IID from some distribution on $\{0,1,2\}$. Conduct a generalized likelihood ratio test to test the null that $X_1, \dots, X_{1000} \overset{\text{IID}}{\sim}$ Binomial$(2,\theta)$ for $\theta \in (0,1)$ against the alternative that the X_i do not follow a binomial distribution. You must report both the generalized likelihood ratio test statistic, the rejection threshold of the test at level $\alpha = 0.05$, and whether or not the test rejects the null at level $\alpha = 0.05$. *Hint: you may use without proof that the MLE of θ is $\hat{\theta} = 179/2000$.*

Solution.

(a) Let O_{jk} denote the entry in the jth row and kth column of the above contingency table. Under H_0, the expected number of entries in each cell is given by

$$E_{jk} = n \times \frac{\sum_{l=1}^{3} O_{jl}}{n} \times \frac{\sum_{r=1}^{2} O_{rk}}{n}.$$

Since $n = 1{,}000$, below is a table of the E_{jk} values:

	$X_i = 0$	$X_i = 1$	$X_i = 2$
Female	352.35	32.805	19.845
Male	517.65	48.195	29.155

Thus the Pearson Chi-squared test statistic for independence is given by

$$T = \sum_{j=1}^{2} \sum_{k=1}^{3} \frac{(O_{jk} - E_{jk})^2}{E_{jk}} \approx 0.905.$$

Under H_0, T approximately follows a chi-squared distribution with $(3-1)\times(2-1) = 2$ degrees of freedom. Hence, according to the table on page A8 in Rice, the rejection threshold for the test at level α is 5.99 (which is the 0.95 quantile of a chi-squared distribution with 2 degrees of freedom). Since $T < 5.99$, we fail to reject the null hypothesis that sex and number of feet with athletes foot are independent.

(b) Let's first prove the hint (this is would have not been necessary; we do this here for so that you understand the background a bit better). We first find the maximum likelihood estimator for θ under the null $X_1, \dots, X_n \overset{\text{IID}}{\sim}$ Binomial$(2,\theta)$. Note that under the null the likelihood is

$$L_0(\theta) = \prod_{i=1}^{n} \binom{2}{X_i} \theta^{X_i}(1-\theta)^{2-X_i} = 2^{\sum_{i=1}^{n} I\{X_i=1\}} \theta^{\sum_{i=1}^{n} X_i} (1-\theta)^{\sum_{i=1}^{n}(2-X_i)} = 2^{81}\theta^{179}(1-\theta)^{1821}.$$

Here, we used that $\binom{2}{1} = 2$ and $\binom{2}{0} = 1$, $\binom{2}{2} = 1$. To find the θ maximizing $L_0(\theta)$, we compute the log-likelihood and take the derivative,

$$l_0(\theta) \equiv \log\left(L_0(\theta)\right) = 81\log(2) + 179\log(\theta) + 1821\log(1-\theta) \Rightarrow l_0'(\theta) = \frac{179}{\theta} - \frac{1821}{1-\theta} \Rightarrow l_0''(\theta) < 0$$

so setting the first derivative of l_0 to 0 and using the 2nd derivative test, it follows that $\hat{\theta} = \frac{179}{2000}$ maximizes the likelihood under H_0.

Now we use the formula on page 342 of Rice for the generalized likelihood ratio test statistic of multinomial cell data. In particular, here

$$-2\log(\Lambda) = 2\sum_{j=0}^{2} O_j \log\left(\frac{O_j}{E_j}\right),$$

where $O_0 = 870$, $O_1 = 81$, $O_2 = 49$ and for $j \in \{0, 1, 2\}$,

$$E_j = n \times \binom{2}{j} \times \hat{\theta}^j (1-\hat{\theta})^{2-j} = \begin{cases} 829.01025 & \text{if } j = 0, \\ 162.9795 & \text{if } j = 1, \\ 8.01025 & \text{if } j = 2. \end{cases}$$

Hence

$$-2\log(\Lambda) = 2\sum_{j=0}^{2} O_j \log\left(\frac{O_j}{E_j}\right) \approx 148.1948.$$

Under H_0, $-2\log(\Lambda)$ approximately follows a chi-squared distribution with 1 degree of freedom. Hence, according to the table on page A8 in Rice, the rejection threshold for the test at level $\alpha = 0.05$ is 3.84 (which is the 0.95 quantile of a chi-squared distribution with 1 degrees of freedom). Since $-2\log(\Lambda) > 3.84$, we reject the null hypothesis that number of feet with athletes foot follows a Binomial$(2, \theta)$ distribution.

5 Problem 5: GLRT and the t-test

Let $X_1, \cdots, X_n$ are i.i.d. $N(\mu, \sigma^2)$, where both μ and σ^2 are unknown. Consider the problem of testing

$$H_0 : \mu = 0$$
$$H_1 : \mu \neq 0$$

Show that the generalized likelihood ratio test statistics for this problem simplifies to

$$\Lambda(X_1, \cdots, X_n) = \left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2 + n\bar{X}^2}\right)^{n/2}$$

Letting $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ and $T = \sqrt{n}\bar{X}/S_X$ (the usual one-sample t-statistics for this problem), show that $\Lambda(X_1, \cdots, X_n)$ is monotonically decreasing function of $|T|$, and hence the generalized likelihood ratio test is equivalent to the two sided t-test which rejects for large values of $|T|$.

Solution. The generalized likelihood ratio test statistics is given by

$$\begin{aligned}
\Lambda(X_1,\cdots,X_n) &= \frac{\sup_{\sigma>0} L(X_1,\cdots,X_n|\mu=0,\sigma)}{\sup_{\mu\neq 0,\sigma>0} L(X_1,\cdots,X_n|\mu,\sigma)} \\
&= \frac{\sup_{\sigma>0}\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp - \sum_{i=1}^n (X_i)^2/2\sigma^2}{\sup_{\mu\neq 0,\sigma>0}\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp - \sum_{i=1}^n (X_i-\mu)^2/2\sigma^2}
\end{aligned}$$

The mle of σ^2 under the null is $\sum_{i=1}^n (X_i)^2/n$ and the mle of μ under alternative is $\bar{X}$ and thus mle of σ under alternative is $\sum_{i=1}^n (X_i-\bar{X})^2/n$ which gives that

$$\begin{aligned}
\Lambda(X_1,\cdots,X_n) &= \frac{\left(\frac{\sum_{i=1}^n (X_i)^2}{n}\right)^{-n/2}\exp(-n/2)}{\left(\frac{\sum_{i=1}^n (X_i-\bar{X})^2}{n}\right)^{-n/2}\exp(-n/2)} \\
&= \left(\frac{\sum_{i=1}^n (X_i-\bar{X})^2}{\sum_{i=1}^n X_i^2}\right)^{n/2} \\
&= \left(\frac{\sum_{i=1}^n (X_i-\bar{X})^2}{\sum_{i=1}^n (X_i-\bar{X})^2 + n\bar{X}^2}\right)^{n/2}
\end{aligned}$$

Now,

$$\begin{aligned}
\Lambda(X_1,\cdots,X_n) &= \left(\frac{\sum_{i=1}^n (X_i-\bar{X})^2 + n\bar{X}^2}{\sum_{i=1}^n (X_i-\bar{X})^2}\right)^{-n/2} \\
&= \left(1 + \frac{n\bar{X}^2}{\sum_{i=1}^n (X_i-\bar{X})^2}\right)^{-n/2}
\end{aligned}$$

Thus, $\Lambda(X_1,\cdots,X_n)$ is a monotonically decreasing function of $n\bar{X}^2/\sum_{i=1}^n (X_i-\bar{X})^2$ and hence monotonically decreasing function of $n\bar{X}^2/S_X^2$ and $|T|$.

6 Problem 6: Unorthodox linear regression

Suppose we observe pairs (X_i, Y_i) which follow the linear regression model without intercept

$$Y_i = \beta X_i + \epsilon_i$$

where we assume $\beta \neq 0$. Typically, in linear regression, we condition on $(X_1,\ldots,X_n)$ and treat them as a constant. In this problem, however, we condition on $(Y_1,\ldots,Y_n)$ and treat them as a constant: this style of analysis has proven very useful in some modern machine learning methods. For simplicity, we assume that $\epsilon_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\sigma^2)$ for known $\sigma^2 > 0$.

(a) Write X_i as a function of Y_i, β, and ϵ_i.

(b) What is the marginal distribution of X_i conditional on $(Y_1, \dots, Y_n)$?

(c) Propose a simple unbiased estimator of $\frac{1}{\beta}$. Argue it is (i) Gaussian and (ii) consistent under the assumption that $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n Y_i^{-2} = \mu < \infty$. *Hint: what is* $\mathbb{E}[X_i]$*?*[2]

(d) Using (c), create an asymptotic $1-\alpha$ confidence interval for $\frac{1}{\beta}$. *Hint: one way to solve this problem is to use a plug-in estimator for the variance and Slutsky's theorem. This may be needed even though* σ^2 *is known.*

(e) Using (d), create an asymptotic $1-\alpha$ confidence interval for β. *Hint: your answer should be very short.*

Solution.

(a) We can write

$$
\begin{aligned}
Y_i = \beta X_i + \epsilon_i &\Leftrightarrow Y_i - \epsilon_i = \beta X_i \\
&\Leftrightarrow \frac{1}{\beta} Y_i - \frac{\epsilon_i}{\beta} = X_i.
\end{aligned}
$$

(b) Since $\epsilon_i/\beta \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2/\beta^2)$, this implies that since we are treating Y_i as a constant,

$$
X_i \overset{\text{ind}}{\sim} \mathcal{N}\left(\frac{1}{\beta} Y_i, \sigma^2/\beta^2\right).
$$

(c) Note that $\mathbb{E}\left[\frac{X_i}{Y_i}\right] = \frac{1}{\beta}$. Thus an unbiased estimator of $\frac{1}{\beta}$ is

$$
\hat{\theta} = \frac{1}{n}\sum_{i=1}^n \frac{X_i}{Y_i}.
$$

$\hat{\theta}$ is Gaussian because it is a linear function of independent Gaussian random variables, since X_i are Gaussian and Y_i are treated as fixed. Furthermore, since the X_i are independent, its variance is

$$
\begin{aligned}
\text{Var}(\hat{\theta}) &= \frac{1}{n^2}\sum_{i=1}^n \frac{\text{Var}(X_i)}{Y_i^2} \\
&= \frac{1}{n^2}\frac{\sigma^2}{\beta^2}\sum_{i=1}^n \frac{1}{Y_i^2} \\
&\to 0
\end{aligned}
$$

where the last step follows because $\frac{1}{n}\sum_{i=1}^n Y_i^{-2} \to \mu < \infty$. Since $\hat{\theta}$ is unbiased and its variance goes to zero, it is consistent.

[2] Note in this problem that we are still treating $\{Y_i\}$ as fixed: it is just some fixed sequence satisfying $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n Y_i^{-2} = \mu$.

(d) We just proved that

$$\frac{n(\hat{\theta} - \frac{1}{\beta})}{\sigma/\beta\sqrt{\sum_{i=1}^n Y_i^{-2}}} \sim \mathcal{N}(0,1).$$

Although we do not know β, we know that $\hat{\theta} \xrightarrow{\text{p}} \frac{1}{\beta}$ and therefore $\frac{1}{\hat{\theta}} \xrightarrow{\text{p}} \beta$ by the continuous mapping theorem. Therefore, by Slutsky's theorem,

$$\frac{n(\hat{\theta} - \frac{1}{\beta})}{\sigma\hat{\theta}^{-1}\sqrt{\sum_{i=1}^n Y_i^{-2}}} \xrightarrow{\text{d}} \mathcal{N}(0,1).$$

As a result, an asymptotic confidence interval for $\frac{1}{\beta}$ is

$$\boxed{\hat{\theta} \pm \Phi^{-1}(1-\alpha/2)\frac{\sigma}{n}\sqrt{\sum_{i=1}^n Y_i^{-2}}.}$$

(e) Let $[L(X), U(X)]$ be the confidence interval for $\frac{1}{\hat{\theta}}$ in the previous problem, so we have that

$$\mathbb{P}(L(X) \leq 1/\beta \leq U(X)) \to 1-\alpha.$$

As a result,

$$\mathbb{P}(1/U(X) \leq \beta \leq 1/L(X)) \to 1-\alpha.$$

This implies that an asymptotic $1-\alpha$ confidence interval for β is $\boxed{\left[\frac{1}{U(X)}, \frac{1}{L(X)}\right]}$.

Distribution on X	Support	PDF or PMF	$\mathbb{E}[X]$	$\mathrm{Var}(X)$
Bernoulli(p)	$\{0,1\}$	$p_X(k) = p^k(1-p)^{1-k}$	p	$p(1-p)$
Binomial(n,p)	$\{0,1,\ldots,n\}$	$p_X(k) = \binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
Poisson(λ)	$\{0,1,2,\ldots\}$	$p_X(k) = \frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ
Geometric(p)	$\{1,2,3,\ldots\}$	$p_X(k) = p(1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Normal(μ,σ^2)	$(-\infty,\infty)$	$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$	μ	σ^2
Exponential(λ)	$[0,\infty)$	$f_X(x) = \lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma(α,β)	$[0,\infty)$	$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
Beta(α,β)	$[0,1]$	$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Table 1: Common distributions table of reference. The third column gives the probability mass function (in the case of discrete random variables) or the probability density function (in the case of continuous random variables). Note that the probability mass functions and probability density functions are defined to be equal to zero outside of the support.