Statistical Theories for Dimensional Analysis

Weijie Shen and Dennis K. J. Lin

The Pennsylvania State University

Supplementary Materials

Supplementary Material A Proofs of the Lemmas and Theorems

Supplementary Material B Data Set of the Phoenix 78 experiment

A. Proofs of the Lemmas and Theorems

Lemma 1. *$(\mathbb{Q}, \mathcal{F})$ is a vector space.*

Proof. Suppose $u = e_1^{\alpha_1} \cdots e_m^{\alpha_m} \in \mathcal{F}$, $v = e_1^{\beta_1} \cdots e_m^{\beta_m} \in \mathcal{F}$, $\lambda \in \mathbb{Q}$. Define addition $u + v = e_1^{\alpha_1+\beta_1} \cdots e_m^{\alpha_m+\beta_m} \in \mathcal{F}$, scaler multiplication $\lambda u = e_1^{\lambda\alpha_1} \cdots e_m^{\lambda\alpha_m} \in \mathcal{F}$. It is easy to verify $\mathcal{F}$ is an Abelian group and $(\mathbb{Q}, \mathcal{F})$ is a vector space. Zero vector is $\mathbf{1} = e_1^0 \cdots e_m^0$. $\square$

Lemma 2. *Suppose unit change T_a transforms fundamental units u_i into $u_i' = a_i u_i$. Then all unit changes $\mathcal{T} = \{T_a : a_i > 0, a_i \in \mathbb{R}, a = (a_1, ..., a_m)^T\}$ form a scaling group. The induced changes on Lebesgue measure $\tilde{\mathcal{T}} = \{\tilde{T}_a : a_i > 0, a_i \in \mathbb{R}, a = (a_1, ..., a_m)^T\}$ is also a scaling group. The induced changes on the measured values of physical quantities $\hat{\mathcal{T}} = \{\hat{T}_a \circ \cdots \circ \hat{T}_a : a_i > 0, a_i \in \mathbb{R}, a = (a_1, ..., a_m)^T\}$ is also a scaling group.*

Proof. Suppose $e_1, \cdots, e_m$ are fundamental dimensions. With measurement system S, dimension e_i has unit u_i. The length of one unit is $\lambda_{u_i}([0, 1]u_i) = \lambda([0, 1]) = 1$. $\lambda([0, 1])$ can be taken as quantity and u_i is the associated unit. Therefore, $\lambda_S([0,1]w) = \lambda([0,1]) = 1$ for all derived units $w = \prod_{i=1}^m u_i^{d_i}$. Let T_a be the mapping (unit change) such that $T_a(S) = S'$, where S' measures dimension e_i by unit $u_i' = a_i u_i$. Then, $\mathcal{T} = \{T_a : a_i > 0, a_i \in \mathbb{R}, a = (a_1, ..., a_m)^T\}$ forms a scaling group: (a) $T_{(a_1,...,a_m)} \circ T_{(b_1,...,b_m)} = T_{(a_1 b_1,...,a_m b_m)}$; (b) $[T_{(a_1,...,a_m)} \circ T_{(b_1,...,b_m)}] \circ T_{(c_1,...,c_m)} = T_{(a_1,...,a_m)} \circ [T_{(b_1,...,b_m)} \circ T_{(c_1,...,c_m)}] = T_{(a_1 b_1 c_1,...,a_m b_m c_m)}$; (c) $T_{(a_1,...,a_m)} \circ T_{(1,...,1)} = T_{(1,...,1)} \circ T_{(a_1,...,a_m)} = T_{(a_1,...,a_m)}$; (d) $T_{(a_1,...,a_m)} \circ T_{(a_1^{-1},...,a_m^{-1})} = T_{(a_1^{-1},...,a_m^{-1})} \circ T_{(a_1,...,a_m)} = T_{(1,...,1)}$.

Plus, we know from physics that the orbit of S under $\mathcal{T}$, $Orb(S) = \{T_a(S) : T_a \in \mathcal{T}\}$, forms $\mathcal{S}$ the collection of all measurement systems. Define $\mathcal{M} = \{\lambda_{S'} : S' \in Orb(S)\}$ to be the collection of physical Lebesgue measures. $\lambda_{S'}([0,1]w) = \lambda_{S'}([0,1] \prod_{i=1}^m a_i^{-d_i} \prod_{i=1}^m u_i'^{d_i})$ $= \prod_{i=1}^m a_i^{-d_i}$. The induced change on Lebesgue measure is $\tilde{T}_a : \mathcal{M} \to \mathcal{M}$ such that $\lambda_{S'} = \tilde{T}_a(\lambda_S)$ and $\tilde{T}_a(\lambda_S)([0,1]w) = \prod_{i=1}^m a_i^{-d_i}$. It can be derived that (a) $\tilde{T}_{a_1 b_1, \cdots, a_m b_m}(\lambda_S) = \tilde{T}_{a_1, \cdots, a_m}(\lambda_S) \times \tilde{T}_{b_1, \cdots, b_m}(\lambda_S)$, as an arithmetic product. The other 3 properties are easy to verify. Therefore, the induced change on Lebesgue measure forms a scaling group $\tilde{\mathcal{T}} = \{\tilde{T}_a : a_i > 0, a_i \in \mathbb{R}, a = (a_1, ..., a_m)^T\}$ by the same structure as $\mathcal{T}$.

Suppose a physical quantity Q has value q and unit w under measurement system S, $q = \lambda_S(Q)$. Then from the above, $\tilde{T}_a(\lambda_S)(Q) = q \prod_{i=1}^m a_i^{-d_i}$. Define $\hat{T}_a : \mathbb{R} \to \mathbb{R}$ such that $\hat{T}_a \circ \lambda_S = \tilde{T}_a(\lambda_S)$. Then $\hat{T}_a(q) = \hat{T}_a(\lambda_S(Q)) = \tilde{T}_a(\lambda_S)(Q) = q \prod_{i=1}^m a_i^{-d_i}$ represents the induced changes on the value q due to unit changes. Let $\hat{\mathcal{T}} = \{\hat{T}_a : a_i > 0, a_i \in \mathbb{R}, a = (a_1, ..., a_m)^T\}$ be the collection of such value transformations. It is straightforward to verify that $\hat{\mathcal{T}}$ is a scaling group. If multiple quantities are of interest, $\hat{\mathcal{T}} = \{\hat{T}_a \circ \cdots \circ \hat{T}_a : a_i > 0, a_i \in \mathbb{R}, a = (a_1, ..., a_m)^T\}$ is still a scaling group. □

Lemma 3. *The probability of an event is dimensionless.*

Proof. Suppose the event A_S is recorded by the measurement system S. According to physics, the characteristic of an event is free from the measurement system. Thus, for any other measurement system S', $P_S(A_S) = P_{S'}(A_{S'})$. Therefore $P(A)$ is dimensionless. $\square$

Lemma 4. *If X is a quantity whose dimension is D and E is a measurable set of values X, then E also has dimension D.*

Proof. Suppose the measurement system is S and its corresponding unit for measuring D is u. Then X has value $x = \lambda_u([0, X])$ with unit u. For any quantity interval $[X_1, X_2]$ (X_1 and X_2 have dimension D), $\lambda_S([X_1, X_2]) = \lambda_u([0, X_2]) - \lambda_u([0, X_1]) = x_2 - x_1$. Suppose in another measurement system S', the unit for D is $u' = au$. Then X has value x/a and $[X_1, X_2]$ has value $(x_2 - x_1)/a$, with unit u'. $\lambda_S([X_1, X_2])/\lambda_S([0, X]) = \lambda_{S'}([X_1, X_2])/\lambda_{S'}([0, X])$. Therefore, interval $[X_1, X_2]$ and $[0, X]$ have the same dimension D. Note that the collection of all intervals with dimension D is a $\pi-$system, and is denoted as I.

Because the domain of X is bounded (in probability), it is easily shown that E and E^c have the same dimension. Suppose $\{E_l, l = 1, 2, \cdots\}$ is a countable sequence of disjoint sets that have the same dimension D. $\lambda_{S'}(E_l) = \lambda_S(E_l)/a$. Therefore, $\lambda_{S'}(\bigcup_l E_l) = \sum_l \lambda_{S'}(E_l) = \sum_l \lambda_S(E_l)/a = \lambda_S(\bigcup_l E_l)/a$. $\bigcup_l E_l$ and E_l have the same dimension D. The collection of all measurable sets with dimension D is a $\lambda-$ system. By Dynkin's $\pi - \lambda$ theorem, all sets within $\sigma(I)$ have dimension D, i.e., any measurable set E for quantity X has dimension D. $\square$

Lemma 5. *If M is the DA transformation that satisfies $M(X_1, \cdots, X_n) = (\pi_{X_{k+1}}, \cdots, \pi_{X_n})^T$, where $\pi_t = X_t X_1^{-b_{t1}} \cdots X_k^{-b_{tk}}$ for $t = k+1, \cdots, n$, then M is maximal invariant over the unit change scaling group $\hat{\mathcal{T}}$ and $(\pi_{X_{k+1}}, \cdots, \pi_{X_n})^T$ is a maximal invariant statistic.*

Proof. Obviously, M is invariant. It is sufficient to prove that, if $M(X_1^{(1)}, \cdots, X_n^{(1)}) = M(X_1^{(2)}, \cdots, X_n^{(2)})$, i.e. $\pi_{X_t^{(1)}} = \pi_{X_t^{(2)}}$ for $t = k+1, \cdots, n$, then there exists a, such that $\hat{T}_a(X_1^{(1)}, \cdots, X_n^{(1)}) = (\hat{T}_a(X_1^{(1)}), \cdots, \hat{T}_a(X_n^{(1)})) = (X_1^{(2)}, \cdots, X_n^{(2)})$. Let $a = (X_1^{(1)}/X_1^{(2)}, \cdots, X_k^{(1)}/X_k^{(2)})$. Then by Lemma 2 $u_i' = T_a(u_i) = X_i^{(1)}/X_i^{(2)} u_i$, $\hat{T}_a(X_i^{(1)}) = X_i^{(1)} \times X_i^{(2)}/X_i^{(1)} = X_i^{(2)}$ for $i =$

$1, \cdots, k$. Since $D_t = D_1^{b_{t1}} \cdots D_k^{b_{tk}}$, $u_t' = T_a(u_t) = \prod_{i=1}^k (X_i^{(1)}/X_i^{(2)})^{b_{ti}} u_t$ for $t = k+1, \cdots, n$. Therefore by Lemma 2, $\hat{T}_a(X_t^{(1)}) = X_t^{(1)} \times \prod_{i=1}^k (X_i^{(2)}/X_i^{(1)})^{b_{ti}} = \pi_{X_t^{(1)}} \prod_{i=1}^k (X_i^{(2)})^{b_{ti}} = \pi_{X_t^{(2)}} \prod_{i=1}^k (X_i^{(2)})^{b_{ti}} = X_t^{(2)}$ for $t = k+1, \cdots, n$. Overall, $\hat{T}_a(X_j^{(1)}) = X_j^{(2)}$ for $j = 1, \cdots, n$. $\square$

Lemma 6. *Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If random vector $(X_1, \cdots, X_n)^T$ follows a continuous distribution F with probability density function f with respect to Lebesgue measure λ, and X_i has dimension D_i. Then $f(X_1, \cdots, X_n)$ has dimension $(\prod_{i=1}^n D_i)^{-1}$, for each given $\omega \in \Omega$.*

Proof. $1 = \mathbb{P}(\Omega) = \int_\Omega d\mathbb{P} = \int_{\mathbb{R}^n} f(x_1, ..., x_n)\lambda(dx_1) \cdots \lambda(dx_n)$. x_i are integral variables with dimension D_i. From definition of Lebesgue integral and Lemma 4, $\lambda(dx_i)$ has dimension D_i. Based on dimensionally homogeneous principle and Lemma 3, $f(x_1, \cdots, x_n)$ has dimension $(\prod_{i=1}^n D_i)^{-1}$. For each $\omega \in \Omega$, $X_i(\omega)$ has dimension D_i. Thus, $f(X_1, ..., X_n)$ has dimension $(\prod_{i=1}^n D_i)^{-1}$, for each given $\omega \in \Omega$. $\square$

Theorem 1. *(Sufficient Dimension Reduction for Parametric Case):*

If Assumptions 1 and 2 hold, θ is dimensionless if and only if $(\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ is a sufficient statistic for θ.

Proof. (Sufficiency)

The joint probability density function of $(Y, X_1, \cdots, X_n)^T$ can be expressed as $p(y, x_1, \cdots, x_n; \theta) = f(y; x_1, \cdots, x_n; \theta)p(x_1, \cdots, x_n)$. From Lemma 3, the conditional probability density function $f(y; X_1, \cdots, X_n; \theta)$ has dimension $(D_0)^{-1}$. Thus, $f(Y; X_1, \cdots, X_n; \theta)Y$ is dimensionless. Since, θ is dimensionless, from Buckingham's Π-Theorem, we have

$$f(Y; X_1, \cdots, X_n; \theta)Y = g(\pi_0; X_1, \cdots, X_k, \pi_{k+1}, \cdots, \pi_n; \theta)$$

$$= g(\pi_0; \pi_{k+1}, \cdots, \pi_n; \theta)$$

The likelihood of θ given $(Y, X_1, \cdots, X_n)^T$ can be expressed as

$$p(Y, X_1, \cdots, X_n; \theta) = f(Y; X_1, \cdots, X_n; \theta) Y p(X_1, \cdots, X_n)/Y$$

$$= g(\pi_0; \pi_{k+1}, \cdots, \pi_n; \theta) p(X_1, \cdots, X_n)/Y$$

By factorization criterion, $(\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ is a sufficient statistic for θ.

(Necessity)

If $(\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ is a sufficient statistic for θ, conditional distribution of $(Y, X_1, \cdots, X_n)^T$ conditioning on $(\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ does not depend on θ. $(\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ are dimensionless, thus changes of measurement system will not change the distribution of $(\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$. If θ is not dimensionless, then changes of measurement system can arbitrarily change the value of θ. Therefore, the distribution of $(\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ does not depend on θ. In summary, the joint distribution of $(Y, X_1, \cdots, X_n)^T$ does not depend on θ, which contradicts with the identifiability of θ. □

Theorem 2. *(Sufficient Dimension Reduction for Nonparametric Case):*

If Assumptions 1 and 2′ hold, a distribution in family $\mathcal{C}$ is invariant to changes in physical dimensions if and only if $T = (\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ is a sufficient statistic for $\mathcal{C}$.

Proof. (Sufficiency)

The joint probability distribution of $(Y, X_1, \cdots, X_n)^T$ is $\mathcal{P} \in \mathcal{C}$ and $\mathcal{C} \ll \lambda$. Suppose $d\mathcal{P}/d\lambda = f_{\mathcal{P}}(y, x_1, \cdots, x_n)$ is the corresponding density function. From Lemma 3, $f_{\mathcal{P}}(y, x_1, \cdots, x_n)$ has dimension $(\prod_{i=0}^{n} D_i)^{-1}$. Thus, $f_{\mathcal{P}}(y, x_1, \cdots, x_n) y x_1 \cdots x_n$ is dimensionless. From Buckingham's Π-Theorem, we have

$$f_{\mathcal{P}}(y, x_1, \cdots, x_n) y x_1 \cdots x_n = g_{\mathcal{P}}(\pi_0, X_1, \cdots, X_k, \pi_{k+1}, \cdots, \pi_n)$$

$$= g_{\mathcal{P}}(\pi_0, \pi_{k+1}, \cdots, \pi_n),$$

where $g_{\mathcal{P}}$ is measurable $\sigma(T)$. Since $\mathcal{P}$ is free from physical dimensions, function $f_{\mathcal{P}}$ and thus $g_{\mathcal{P}}$ will not change with dimensions. The arguments of $g_{\mathcal{P}}$ are dimensionless. Therefore the value of $g_{\mathcal{P}}(\pi_0, \pi_{k+1}, \cdots, \pi_n)$ does not change with dimensions. Then $d\mathcal{P}/d\lambda = g_{\mathcal{P}}(\pi_0, \pi_{k+1}, \cdots, \pi_n)/yx_1 \cdots x_n$. By factorization criterion, $T = (\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ is a sufficient statistic for $\mathcal{C}$.

(Necessity)

Assume $(\pi_0, \pi_{k+1}, \cdots, \pi_n)^T$ is a sufficient statistic for $\mathcal{C}$. From the definition, $\forall A \in \mathcal{R}^{n+1}, \exists \kappa_A$ measurable $\sigma(T) \subseteq \mathcal{R}^{n+1}$, such that $\mathcal{P}(A|T) = \kappa_A$ a.s. $\mathcal{P}, \forall \mathcal{P} \in \mathcal{C}$. Thus, $\mathcal{P}(A) = \int_{\mathbb{R}^{n+1}} \kappa_A d\mathcal{P} = \int_{\mathbb{R}} \kappa_A \circ T^{-1} d\mathcal{P}_T$. T is free from physical dimensions, thus its distribution $\mathcal{P}_T$ should be free from dimensions. That is to say, suppose the changes in dimensions turn $\mathcal{P}$ into $\mathcal{P}'$, then $\mathcal{P}_T = \mathcal{P}'_T$. This leads to $\mathcal{P}(A) = \mathcal{P}'(A), \forall A \in \mathcal{R}^{n+1}$, i.e. $\mathcal{P} = \mathcal{P}'$. In summary, as long as $\mathcal{C}$ is identifiable, $\mathcal{P}$ is invariant to changes in physical dimensions, for $\forall \mathcal{P} \in \mathcal{C}$. □

Theorem 3. *(Completeness)*

If Assumptions 1 and 3 hold, $(\pi_0, \pi_{k+1}, \cdots, \pi_n)$ *is complete for family* $\mathcal{C}$,

$$\forall F \in \mathcal{C}, E_F h(\pi_0, \pi_{k+1}, \cdots, \pi_n) = 0 \Rightarrow \forall F \in \mathcal{C}, \mathbb{P}_F(h(\pi_0, \pi_{k+1}, \cdots, \pi_n) = 0) = 1.$$

Proof. If $\exists F \in \mathcal{C}$ s.t. $\mathbb{P}_F(h(\pi_0, \pi_{k+1}, \cdots, \pi_n) = 0) < 1$, then $\mathbb{P}_F(h(\pi_0, \pi_{k+1}, \cdots, \pi_n) \neq 0) >$

0. Arbitrarily suppose $\mathbb{P}_F(h(\pi_0, \pi_{k+1}, \cdots, \pi_n) > 0) > 0$. Then the conditional distribution $(\pi_0, \pi_{k+1}, \cdots, \pi_n)|(h(\pi_0, \pi_{k+1}, \cdots, \pi_n) > 0) \sim F'$ can be properly defined and its parameters are also dimensionless. $F' \in \mathcal{C}$ and apparently $E_{F'}h(\pi_0, \pi_{k+1}, \cdots, \pi_n) = E_F[h(\pi_0, \pi_{k+1}, \cdots, \pi_n)$ $|h(\pi_0, \pi_{k+1}, \cdots, \pi_n) > 0] > 0$, which contradicts with the condition. $\square$

B. Data Set of the Phoenix 78 experiment

Table B.1: Data Set of the Phoenix 78 experiment

DATE	w	z	w_*	z_i	$\pi_0 = w/w_*^2$	$\pi_1 = z/z_i$
90978	1.082	145.2	2.070	1452	0.252	0.100
90978	1.115	160.7	1.973	1236	0.286	0.130
90978	1.388	161.1	1.952	1151	0.364	0.140
90978	0.764	196.9	1.245	547	0.493	0.360
90978	0.544	165.3	1.120	435	0.434	0.380
90978	1.747	418.9	1.944	1074	0.462	0.390
90978	1.041	420.8	1.926	1002	0.281	0.420
90978	0.616	250.6	1.230	522	0.407	0.480
90978	0.713	262.2	1.238	535	0.465	0.490
90978	0.469	298.9	1.083	421	0.400	0.710
90978	1.223	762.5	1.944	1074	0.324	0.710
90978	0.628	771.4	1.930	1015	0.169	0.760
90978	0.394	397.8	1.225	510	0.262	0.780
90978	0.387	396.0	1.219	495	0.260	0.800
90978	0.114	388.1	1.014	396	0.111	0.980
90978	0.181	408.0	1.047	408	0.166	1.000
90978	0.191	575.3	1.257	564	0.121	1.020
90978	0.105	601.7	1.245	547	0.068	1.100
92178	1.301	152.2	1.668	1072	0.467	0.142
92178	1.373	151.5	1.681	1052	0.486	0.144
92178	1.640	151.8	1.816	748	0.497	0.203
92178	1.385	151.7	1.757	712	0.449	0.213
92178	1.694	552.7	1.695	1037	0.590	0.533
92178	1.774	552.9	1.732	1022	0.591	0.541
92178	0.760	403.2	1.700	672	0.263	0.600
92178	0.913	403.2	1.650	640	0.335	0.630
92178	0.988	863.0	1.668	1072	0.355	0.805

Table B.2: Data Set of the Phoenix 78 experiment (Continued table)

DATE	w	z	w_*	z_i	w/w_*^2	z/z_i
92178	0.719	863.2	1.680	1054	0.255	0.819
92178	0.241	612.9	1.605	608	0.094	1.008
92178	0.226	612.9	1.564	567	0.092	1.081
92178	0.407	1163.5	1.696	1037	0.141	1.122
92178	0.349	1163.3	1.728	1024	0.117	1.136
92278	1.123	160.3	1.706	1394	0.386	0.115
92278	1.302	160.1	1.896	1291	0.362	0.124
92278	1.453	249.7	1.882	1255	0.410	0.199
92278	1.341	560.6	1.869	1224	0.384	0.458
92278	1.312	709.7	1.771	1378	0.418	0.515
92278	1.016	859.6	1.855	1168	0.295	0.736
92278	0.654	1170.7	1.818	1366	0.198	0.857
92278	0.407	1100.5	1.841	1095	0.120	1.005
92278	0.102	1619.4	1.856	1354	0.030	1.196
92778	1.578	150.0	1.652	990	0.578	0.152
92778	1.908	329.9	2.135	1907	0.419	0.173
92778	0.861	150.0	1.568	534	0.350	0.281
92778	0.760	150.0	1.344	451	0.421	0.333
92778	1.631	400.0	1.600	923	0.637	0.433
92778	0.489	220.0	1.302	422	0.288	0.521
92778	0.687	280.0	1.531	518	0.293	0.541
92778	2.736	940.1	2.025	1660	0.667	0.566
92778	1.325	609.1	1.573	870	0.536	0.700
92778	0.321	280.0	1.263	394	0.201	0.711
92778	0.438	370.0	1.491	502	0.197	0.737
92778	2.539	1090.0	1.914	1420	0.693	0.768
92778	0.634	850.0	1.557	822	0.262	1.034
92778	1.764	1220.0	1.827	1170	0.528	1.043
92778	0.060	380.0	1.219	363	0.040	1.047
92778	0.129	520.0	1.451	488	0.061	1.066
92778	0.273	1010.0	1.542	773	0.115	1.307